ภาคผนวก

การหาค่าเบี่ยงเบนมาตรฐานของคิวแบบ Geom/Geom/1

จากหัวข้อ 6.4.1 $\pi_1 = \frac{p(1-\rho)}{q}$

และ ค่าเฉลี่ยของคิว Geom/Geom/1 หรือ $E(Q) = \frac{p}{q\ (1-\rho)}$

$Var = E(Q^2)^* - E(Q)^2$

$$= \left(\frac{1+\rho}{(1-\rho)^3}\frac{p(1-\rho)}{q}\right) - \frac{p^2}{q^2(1-\rho)^2}$$

$$= \frac{p(1-\rho)}{q(1-\rho)^2} - \frac{p^2}{q^2(1-\rho)^2}$$

$$= \frac{pq+\rho pq-p^2}{q^2(1-\rho)^2}$$ เนื่องจาก $q = \frac{1}{1+\rho}$ เมื่อ p=0.5 ดังนั้น

$$= \frac{p-p^2}{q^2(1-\rho)^2}$$

$$= \frac{p(1-p)}{q^2(1-\rho)^2}$$

Standard deviation (std) $= \sqrt{E(Q^2)-E(Q)^2}$

$$= \sqrt{\frac{p(1-p)}{q^2(1-\rho)^2}}$$

$$= \frac{p}{q\ (1-\rho)}$$ เมื่อ p=0.5

*วิธีหา $E(Q^2)$

$$E(Q) = \sum_{i>=1} i\pi_i = \pi_1 + 2\pi_2 + 3\pi_3 + \ldots\ldots \quad ; \ \pi_i = \rho^{i-1}\pi_1$$

$$E(Q) = \sum_{i>=1} i\pi_i = (1 + 2\rho + 3\rho^2 + 4\rho^3 + \ldots\ldots)\pi_1$$

$$E(Q^2) = \sum_{i>=1} i^2\pi_i = \pi_1 + 4\pi_2 + 9\pi_3 + \ldots\ldots \quad ; \ \pi_i = \rho^{i-1}\pi_1$$

$$E(Q^2) = \sum_{i>=1} i\pi_i = (1 + 2\rho + 3\rho^2 + 4\rho^3 + \ldots\ldots)\pi_1$$

พิจารณาอนุกรม $\sum_{i=0}^{\infty} \rho^i = (1 + \rho + \rho^2 + \rho^3 + \ldots\ldots)$ เป็นอนุกรมเรขาคณิต

และ $\sum_{m=0}^{\infty} \rho^m = (1 + \rho + \rho^2 + \rho^3 + \ldots\ldots)$ เป็นอนุกรมเรขาคณิต

ดังนั้น $\lim_{n\to\infty} \sum_{n=0}^{\infty} \rho^n = \lim_{n\to\infty} (1 + \rho + \rho^2 + \rho^3 + \ldots\ldots + \rho^n) = \frac{1}{1-\rho}$, $|\rho| < 1$

$$\sum_{i=0}^{\infty} \rho^i \sum_{m=0}^{\infty} \rho^m = (1 + \rho + \rho^2 + \rho^3 + \ldots\ldots)\ (1 + \rho + \rho^2 + \rho^3 + \ldots\ldots)$$

$$= 1 + 2\rho + 3\rho^2 + 4\rho^3 + \ldots\ldots$$

$$= \left(\frac{1}{1-\rho}\right)\left(\frac{1}{1-\rho}\right)$$

$$= \frac{1}{(1-\rho)^2}$$

ดังนั้น $E(Q) = \frac{1}{(1-\rho)^2}\ \pi_1$

ให้ a $= 1 + 2\rho + 3\rho^2 + 4\rho^3 + \ldots\ldots + n\rho^{n-1}$

b $= 1 + 4\rho + 9\rho^2 + 16\rho^3 + \ldots\ldots + n^2\rho^{n-1}$

a' $= 2 + 6\rho + 12\rho^2 + \ldots\ldots + (n^2 - n)\rho^{n-2}$

a+b $= 2 + 6\rho + 12\rho^2 + \ldots\ldots + (n^2 + n)\rho^{n-1}$

จะได้ว่า $a+b = a' + (n^2 + n)\rho^{n-1}$

พิจารณา $\lim_{n\to\infty} \sum (n^2 + n)\rho^{n-1} = 0$ เมื่อ $|\rho| < 1$

ดังนั้น $b = a' - a$

เนื่องจาก $a = \frac{1}{(1-\rho)^2}$ ดังนั้น $a' = \frac{2}{(1-\rho)^3}$

$$b = \frac{2}{(1-\rho)^3} - \frac{1}{(1-\rho)^2}$$

$$= \frac{1+\rho}{(1-\rho)^3}$$ ดังนั้น $E(Q^2) = \frac{1+\rho}{(1-\rho)^3}\ \pi_1$ □

บทความของผู้วิจัย

1. Chiewthanakul B. and Waiyanon, K. 2006. The Performance on the Prevention of Packet Reordering in load-balanced router. Proceedings of the Technology and Innovation for Sustainable Development Conference (TISD). Khon Kaen, Thailand.199-208.

2. กนกวรรณ ไวยนนท์และพิเชษฐ เชี่ยวธนะกุล., ประสิทธิภาพการป้องกันการจัดเรียงตัวใหม่ของแพ็กเก็ตในเราเตอร์แบบถ่วงสมดุลย์.วารสารวิจัย มข. (ฉบับบัณฑิตศึกษา) ก.ค.-ธ.ค. 2549;ปีที่ 6(2).

The Performance on the Prevention of Packet Reordering in Load-Balanced Router

Chiewthanakul Bhichate and Kanokwan Waiyanon
Department of Computer Engineering, Khon Kaen University
Khon Kaen 40002, Thailand
{bhichi, kansoi}@kku.ac.th

Abstract- **In this paper, building on the two-stage switch, we proposed an algorithm called Adaptation-Size-Frame-RR with Three Dimensional Queues (ASF-RR with 3DQs) that can be prevent the problem of packets out of sequence in the load-balanced router in which different packets of the same flow may be taken into different paths. Our model of the problem is derived from a problem present in two-stage load-balanced switches, which have first mesh fabric between input buffer stage and intermediate buffer stage and the second one between intermediate buffer stage and output buffer stage. The packets are split into cells in which are grouped into flows in Virtual Output Queues at these buffer stages. The cells of each flow are balanced over center switched elements. We use the frame-based scheme called round-robin order to pick a fixed size packet or cell from none empty Virtual Output Queues in first-stage and spread it like a Uniform Frame Spreading. For this algorithm, we show that it prevents miss-sequencing while maintaining the high performance benefits in terms of throughput and delay of the basic-two stage switch.**

Keywords: Load-Balanced Router, Virtual Output Queues, Packet Reordering

I. INTRODUCTION

A. From Scheduling to Load-Balancing Routing

CURRENT Internet core routers commonly implement combined input and output queueing (CIOQ) with a centralized scheduler. Numerous centralized scheduling algorithms have been proposed in the literature [1], [2], [3]. Nevertheless, although these scheduling algorithms can theoretically provide a guaranteed throughput of 50% to 100% ([4], [5], [6]), they are becoming impractical as the line rates and numbers of ports grow, because of their complexity and/or the speedup of buffer memory

There has been recent interest in a new approach, which eliminates scheduling, using a load-balance switch architecture [7], [8], [9], [10], [11]. This architecture is based on load-balancing packets uniformly inside the router before forwarding them to their correct destination, an idea first introduced by Valiant [12]. C.S. Chang *et al.* [10] show that a load-balanced router by using VOQ (Virtual output queue) does not require any scheduler and that it can guarantee 100% throughput for a broad class of traffic. Therefore, the load-balanced router is not subject to the two main problems commonly present in former architectures: centralized scheduling and the lack of throughput guarantees needed by network operators. This makes the load-balanced router an appealing architecture to study. However, the load-balanced router suffers from several problems, such as packet reordering, which defined as packets from a same flow depart from the router in an order different from the one in which they arrived, and the need for frequent switch fabric reconfigurations. Consequently, the TCP (Transmission control protocol) does not perform well when out-of-order packets arrive at their destination. Out-of-order packets can be perceived as loss indicators and trigger unnecessary retransmissions and TCP timeouts [13]. These retransmissions and timeouts cause a decrease in TCP throughput and an increase in packet delay. Therefore, since TCP traffic constitutes the vast majority of Internet traffic [14], network operators generally insist that routers do not reorder packets belonging to the same application flow.

Figure 1 illustrates the possibility of the packet reordering in a load-balanced router. In this example, all packets considered are destined to the output 1. By definition, in the first N time-slots, each input is connected to the first

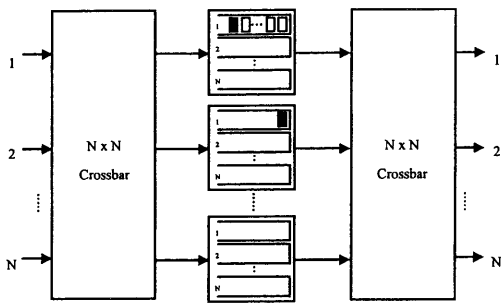


Fig. 1. Example of packet reordering in a load-balanced router.

intermediate input exactly once. Assume that each input except the first one receives a packet and transfers it to the first intermediate input when they are connected together. After N time-slots, the first intermediate input will have received $N-1$ packets from inputs $2,\ldots,N$, and only one of these packets will have been serviced. Therefore, $N-2$ packets are left in the first intermediate input, which are represented as transparent. Assume that the first input then receives two packets back-to-back when it is successively connected to intermediate inputs 1 and 2. Therefore, the first input consecutively transfers these two packets to intermediate inputs 1 and 2, which are represented as filled. The second packet is alone in its queue, while the first one has $N-2$ packets in front of it. Therefore, the second packet is serviced earlier. It arrives at output 1 earlier, and leaves the router earlier. Consequently, the two packets are reordered. There are two methods of preventing packet reordering. The first method consists in bounding the amount of reordering and using a finite reordering buffer at the output. The second method is to make sure that packets arrive in order to the output, thus keeping packets in order throughout the router. We saw in the example above that reordering can occur when VOQ lengths are different, and that the amount of reordering typically increases as the difference of VOQ lengths increases. Therefore, most of the algorithms that prevent reordering will try to bound or prevent any VOQ length difference. In [15], the authors propose two based on the first method of bounding reordering. Both schemes rely on algorithms found in the parallel packet switch router and use small input-stage coordination buffer. The first scheme, called FCFS (First Come First Served), uses a jitter control buffer in each intermediate input to ensure proper ordering of the traffic leaving the intermediate inputs. The second scheme, EDF (Earliest Deadline First), schedules packets according to their departure times in an ideal router. However, both schemes do not seem practical. The jitter control mechanism in FCFS might require up to N memory-write accesses per time slot. And EDF needs to retrieve the packet with the smallest time stamp from a queue, making it hard to implement in a high performance router. The second method of preventing reordering, which keeps packets ordered throughout the router, is used in [10], [11] and [16]. In [10], the authors introduce an algorithm in which the buffers in the intermediate inputs are finite and packets are guaranteed to leave the router in order. This algorithm assumes that in each frame period, arrivals of packets destined to a given output are constrained. However, this property is not satisfied in general in the Internet. In [11], the authors present an algorithm that uses a coordination buffer in the input stage, and then queues packets in VOQs in the intermediate inputs based on their input, intermediate input, and output. Using this fine-grained queueing, the algorithm guarantees that packets arrive at the outputs in order. However, this algorithm requires a more complex queueing management system, and communication of state information between the intermediate inputs and the output. In [16], the authors propose the scheme, called FOFF (Full Ordered Frames First), allows some bound reordering inside the router, and relies on a reordering buffer at the output. Since this control mechanism allows for a non-empty queue to send packets when no queue has any full frame. Therefore, FOFF avoids starvation but allows for different VOQ sizes and increasing the miss-sequence packets store up in the VOQs.

Since, all these algorithms present significant problems and are not satisfied in the Internet. Therefore, our objective is to find a different scheme in order to guarantee that packets leave the router in order and to prevent miss-sequencing and bound the miss-sequence packets store up in the VOQs. This scheme also provides throughput and delay guarantees for all traffic patterns. The paper is organized as follows. We first introduce the switch architecture in Section II. The details of the Adaptation-Size Frames-RR with 3DQ mechanism will be presented in Section III. In this section, we will show how this mechanism

helps prevent miss-sequencing and bound the miss-sequence packets in the VOQs. Then we will prove some theorems on its delay and throughput in Section IV. Finally, the conclusion is given in Section V.

II. SWITCH ARCHITECTURE

A. Assumptions

Throughout this paper, we assume for simplicity that all incoming variable-size packets are segmented into fixed-size packets or cells and reassembled when leaving the router. We say that links are connected to routers through linecards, and denoted by N the number of input (and output) linecards. We also assume that all linecards have the same line rate and that time is slotted, so that at most one cell can arrive at any input port and at most one cell can depart from any output port in each time-slot. Finally, we assume that initially there is no packet in the router.

B. Load-balanced Router Architecture

The switch architecture that we will propose as the basis for this paper is developed from [11], and shown in Figure 2. The load-balanced router architecture relies on two fixed equal-rate switches and three buffer stages. The buffers at the first stage help to limit the amount of miss-sequencing. The first switch is used to prevent the head-of-line blocking by distribute the packets in a round-robin fashion to the second stage buffers or VOQs. Thus, the first switch performs load-balancing for incoming traffic. The packet will then be transferred across the second switch to third stage buffers. These last stage buffers are used for reassembly of a packet when leaving the router. Each of the two switching stages goes through the same pre-determined cyclic shift configuration. That is, at time t, input i of each switching stage is connected to output $[(i+t-1) \bmod N]+1$. Therefore, each of these switches connects any input to any output exactly $1/N$-th of the time, regardless of the arrival traffic. These fixed equal rate switches could be implemented using N x N switch fabrics that are reconfigured at every time-slot. However, if we were to use an optical switch fabric, this constraint would hinder our ability to scale to higher speeds. In addition, a system with no reconfigurations is obviously simpler and more reliable than a system with frequent reconfigurations. Therefore, we would like to replace these fixed equal-rate switches with a fixed mesh of optical channels. A first observation is that we can replace each fixed equal-rate switch with N^2 fixed channels at rate R/N, as illustrated in Figure 2. Hence, a uniform mesh replaces each switch. The rate provided between any switch input and any switch output will stay the same. Therefore, the switch is still a fixed, equal-rate switch. A second observation is that we can replace the two meshes with a single mesh running twice as fast, as shown in Figure 3a). Since, every cell traverses the mesh twice, each time at rate R/N; therefore, the total channel rate is $2R/N$. After a cell crosses the switch the first time, it

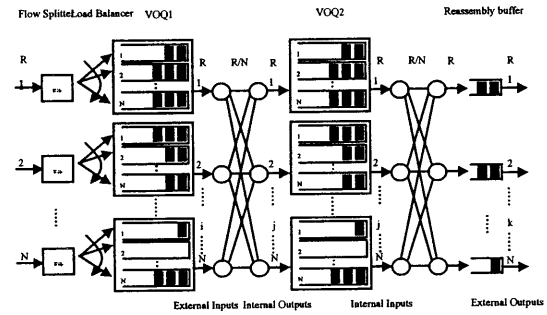


Fig. 2. Load-balanced router architecture based on a double mesh.

is stored in an intermediate linecard; from there, it crosses the switch again to reach the output linecard. A third observation is that a uniform mesh with optical channels can be replaced an Arrayed Waveguide Grating Router (AWGR), as represented in Figure 3b).

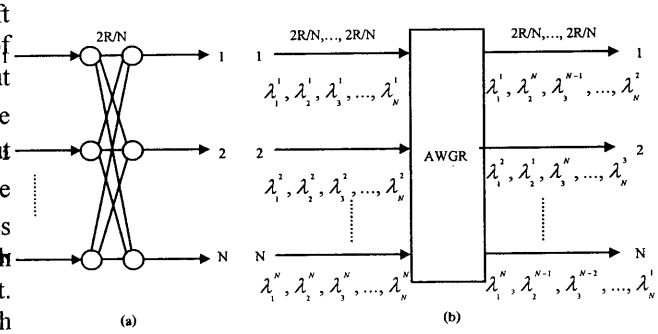


Fig. 3. Load-balanced router architecture based on (a) a single mesh, and (b) an AWGR.

Input i transmits N district channels on its outgoing fiber. Each different channel λ_w^i is

transmitted at rate $2R/N$ on specific wavelength λ_w. The AWGR, a passive optical device, shuffles the channels such that each channel of a given input is destined to a different output. As a result, the system behaves as a single mesh. Furthermore, its main advantage is the reduction in the number of fibers needed from N^2 to $2N$.

C. Definitions

The switch architecture consists of two stages of switching. The inputs of the first stage are called external inputs (EIs), and numbered $i = 1,\ldots,N$. The outputs of the first stage, called internal outputs (IOs), are collocated with the inputs of the second stage, called internal inputs (IIs). IOs and IIs will be used interchangeably in this paper, and are numbered $j = 1,\ldots,N$. Finally, the outputs of the second stage, called external output (EOs), are numbered $k = 1,\ldots,N$.

Let's follow the path of packets through the switch.

1. First, a flow splitter labels each packet in EI i as belonging to a given flow F_{ik}, where k is the EO to which this packet is destined. There are therefore N possible flows per EI representing the N different EOs to which the packets may be destined.
2. A load balancer sends all the packets from F_{ik} to the N VOQ1s, in a round-robin manner –i.e. the first packet from a given flow is sent to the VOQ1 for IO 1, the second one is sent to the VOQ1 for IO 2, and so on, independently of the packet arrival times. If the flow has not enough packets to fill up for N VOQ1s, then the null packet that has same size will be added at the remain of the VOQ1s. Because the load balancers are not necessarily synchronized with the sequence of configurations of the first-stage switch, arriving packets are buffered and do not necessarily immediately leave the VOQ1s. Note that the queueing length in each VOQ1 has maintained the same size to limit the amount of the unbounding miss-sequence packets that will be stored up in the VOQ2s.
3. The VOQ1s are served in deterministic order by the first stage switch, and when their turn comes the packets leave their VOQ1 and pass through the first-stage switch.
4. After leaving the first-stage switch the packets are queued in the VOQ2s. The inputs of the VOQ2s are the IIs, and their outputs are the EOs.
5. The VOQ2s are served in deterministic order by the second-stage switch, and when their turn comes the packets leave their VOQ2 and pass through the second-stage switch.
6. After leaving the second-stage switch the packets are rearranged in the reassembly buffers so that all of the queues containing cells from a given external input are adjacent to each other.
7. Finally, the packets leave the reassembly buffer and exit through the EO.

The following property of the switch will prove useful in this paper (proved in [15]).

Property 1 If a cell arrives to the switch at time t, it will arrive to the VOQ2s no sooner than t, and no later than $t + N^2$.

D. EDF-3DQ, An extension of VOQ

The return of HOL blocking can be occurred if we use the typical VOQ2. Consider a packet, p, that sits in the VOQ2(j,k). We'll assume that p was the earliest arriving packet to the switch among all packets in its VOQ2(j,k), but that p is not currently sitting at head-of-line (HOL) in its VOQ2. Packet p is obviously the earliest arriving packet of its flow F_{ik} in VOQ2(j,k), and therefore shall be sat in front of the other packets of its flow in VOQ2(j,k). However, this packet is blocked by packets ahead of it that arrived later to different external inputs and also scheduled to depart from EO k. This is classical HOL blocking, and the solution is to subdivide each VOQ2 into a separate queue for each external input.

Now, we consider typical VOQ2s that transform one-dimensional queues into two-dimensional queues, one per (input, output) pair. There are therefore N^2 VOQ2s. In this switch, we will use three-dimensional queues

(3DQs), with a different queue per (i, j, k); hence, there are nows a total N^3 3DQs. From hereon, we'll assume that we replace the VOQ2s by 3DQs. With 3DQs, the earliest cell for (j, k) is always the HOL cell in its queue. Suppose that the earliest deadline first (EDF) algorithm is used with 3DQ structure to bound miss-sequencing by serving cells in the VOQ2s in the order that they arrived to the switch, rather than strictly from the head of line. We only need a comparison among N timestamps. However, this simplification comes at the cost of using N^3 3DQ instead N^2 VOQ2s. The EDF has the following properties, proved in [15].

Property 2 Packet miss-sequencing is bounded by $2N^2 + N$.

Note that it is therefore possible to add a finite resequencing buffer after the switch for each external output.

Property 3 The Packet delay in EDF is bounded by the sum of the packet delay in a first-come-first-served (FCFS) OQ switch, and a constant equal to $2N^2 + N$.

III. ADAPTATION-SIZE FRAMES-RR WITH 3DQ MECHANISM

A. Background

Supposed that the algorithm deals with frames, where one frame consists of N packets. First, the algorithm called Adaptation-Size Frames-RR (ASF-RR) will add null packets to the non-full frames and then service them in round-robin order. Note that because the average delay is independent of the order in which the frames are serviced and for heavy traffic environment (i.e., incoming traffics can be filled on full frames without null packets), ASF-RR will have the same average delay as OQ-FCFS (the FCFS version of output queue). If there is at least one full frame in the queue, then ASF-RR is not idle. Therefore, the ASF-RR is work-conserving for full frames, in the sense that if there is at least one full frame left, then there is at least one full frame being serviced.

B. A Combination of ASF-RR and 3DQs

We apply ASF-RR to the 3DQs in the second-stage of the switch. To understand its operation, consider external output k. We'll define *cycle* to be the set of N consecutive time slots during which EO k receives cells successively from IIs 1 through N, and we'll define the *candidate set* of 3DQs for (i, k) as $\{(i,1,k),(i,2,k),\ldots,(i,N,k)\}$.

Assume that the last serviced cell in the candidate set came from II j_{last}. Then, because of the properties of the load-balancer, we know that the next in-order cell for the flow F_{ik} will come from II $j_{last} + 1 \pmod N$. Let p_{ik} be the pointer to the II of the next in-order cell: $p_{ik} \leftarrow j_{last} + 1 \pmod N$

We define the *frame* for (i, k) as $f(i,k) = \{(i, p_{ik}, k),(i, p_{ik}+1, k),\ldots,(i, N, k)\}$, and we will say that frame $f(i, k)$ is *full* if every 3DQ (i, j, k) for $j = \{p_{ik}, \ldots, N\}$ is non-empty. We can see that if the frame is full then its next in-order cell is in 3DQ (i, p_{ik}, k), the one after is in $(i, p_{ik}+1, k)$, and so on, up until (i, N, k). Suppose that the earliest deadline first (EDF) algorithm is used with 3DQs structure. Therefore, at the beginning of every cycle, the 3DQs have been rearranged so that all of the queues containing cells from a given external input are adjacent to each other. This is the key to preventing and bounding the cells within a frame from becoming miss-sequenced.

In ASF-RR with 3DQs mechanism, an external output reads all the cells in a full frame from an external input, before moving on to read a full frame from next external input. External output k uses round-robin pointer $p_{ff}(k)$ to remember which EI the last full frame came from.

More precisely, the following computation is performed by external output k at the beginning of every cycle: Starting at $p_{ff}(k)$, find the next frame arrived from EI i_{ff}, then update $p_{ff}(k) \leftarrow i_{ff} + 1 \pmod N$.

IV. ASF-RR WITH 3DQS PERFORMANCE

In this section we show that the average delay for ASF-RR with 3DQs algorithm is less than the average delay for OQ plus a constant and that ASF-RR with 3DQs has the same throughput as OQ. The proofs is rely on the

observation that ASF-RR with 3DQs is work-conserving for full frames.

A. Definitions

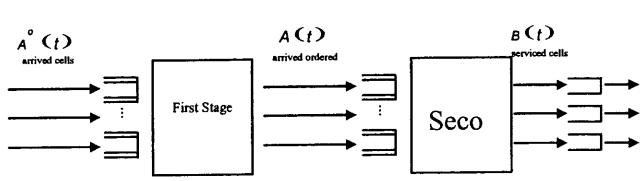
For simplicity, we will only consider the cells destined to a given EO k. We'll define the following values, as illustrated in Figure 4.

1. $A_i^\circ(t)$ is the cumulative number of cells destined to EO k that have arrived to EI i up to and including time-slot t. It is therefore the index of the last cell from F_{ik} that has arrived to EI i.

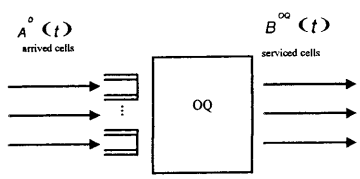
2. $A^\circ(t) = \sum_{i=1}^{N} A_i^\circ(t)$ is the total number of cells destined to k that have arrived to the switch up until t.

3. $A_i(t)$ is the index of the last cell in flow F_{ik} to have arrived to an II by the beginning of the current cycle, without any cells missing from the flow's FCFS order, without any cells missing in front of them, the *ordered cells*. The total number of ordered cells is $A(t) = \sum_{i=1}^{N} A_i(t)$. By definition, if t' is the beginning of a cycle for EO k (that is, $t' = k + 1 + l \cdot N$, with l an integer), then $A_i(t) = A_i(t')$ when $t' \le t \le t' + N - 1$. In addition, as shown in Property 1, $A_i^\circ(t' - N^2) \le A_i(t') \le A_i^\circ(t')$ (assume that these values are zero for $t' \le 0$).

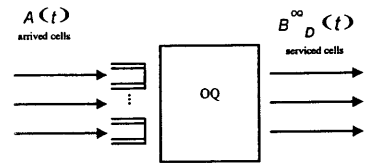
4. $B_i(t)$ is the number of ordered cells that have already been served by the second stage up until time-slot t, and $B(t) = \sum_{i=1}^{N} B_i(t)$.



(a) Two-stage switch with AFS-RR with 3DQ



(b) OQ model for first stage (with $A^\circ(t)$ arrivals)



(c) OQ model for second stage (with $A(t)$ arrivals)

Fig. 4. Illustration of terminology

5. $q(t) = A(t) - B(t)$ is the number of ordered cells queued in the 3DQs that are destined to k

6. $A_{ff}(t) = \sum_{i=1}^{N} N \cdot \left\lceil \frac{A_i(t)}{N} \right\rceil$ is the number of cells in full frames already arrived to the 3DQs and destined to k.

7. Likewise, $B_{ff}(t) = \sum_{i=1}^{N} N \cdot \left\lfloor \frac{B_i(t)}{N} \right\rfloor$ is the number of cells in full frames that have already been served by the second stage.

8. $ff(t) = \frac{A_{ff}(t)}{N} - \frac{B_{ff}(t)}{N} = \sum_{i=1}^{N} \left\lceil \frac{A_i(t)}{N} \right\rceil - \left\lfloor \frac{B_i(t)}{N} \right\rfloor$ is the number of full frames queued in the 3DQs, and $q_{ff}(t) = N \cdot ff(t) = A_{ff}(t) - B_{ff}(t)$

9. $B^{OQ}(t)$ and $q^{OQ}(t)$ are respectively the cumulative number of services and the length of the queue in and FCFS OQ switch of speed-up 1, where the cumulative number of arrivals is $A^\circ(t)$.

10. Similarly, $B_D^{OQ}(t)$ and $q_D^{OQ}(t)$ correspond to what we will call the delayed version of OQ, which is computed with a cumulative number of arrivals equal to $A(t)$. In other words, it represents an OQ switch of speed-up 1 with the same ordered arrivals as seen by the IIs in the two-stage switch. There are some well-known properties of OQ apply:

$$q_D^{OQ}(t) = A(t) - B_D^{OQ}(t),$$

$$q_D^{OQ}(t) = \max_{0 \le s \le t} (A(t) - A(s) - (t-s)),$$

$$B_D^{OQ}(t) = \max_{0 \le s \le t} (A(s) + (t-s)), \text{ and}$$

$B_D^{OQ}(t) - B_D^{OQ}(s) \le t - s$, whenever $s \le t$ [11].

11. $t' = k + 1 + l \cdot N$, where l is an integer, is any time-slot when the cycle for EO k begins.

B. ASF-RR with 3DQs Average Delay within a constant from OQ

In this section we will show that the average delay for ASF-RR with 3DQs is within a constant delay of the average delay for an OQ switch for the same arriving traffic.

We will first compare ASF-RR with 3DQs with the *delayed OQ*, which is an OQ having the same ordered arrivals as the second stage. We will show that ASF-RR with 3DQs is work-conserving for full frames and as many cells as the delayed OQ model, with a queue size almost as small. This results in a bounded average delay difference with the delayed OQ model. Then we compare the delayed OQ model with a regular OQ switch having the same packet arrivals as the first stage. Using a delay bounded average delay difference between ASF-RR with 3DQs and an OQ switch.

We start by establishing that whenever there is at least one frame arrival to 3DQs, the number of serviced frames increases by one in the next cycle.

Lemma 1 If at least one frame arrives to the second stage, then $B_{ff}(t' + N) = B_{ff}(t') + N$.

Proof: Since every frame arrives to the second stage is always full frames. Therefore, $ff(t') > 0$ and by the property of 3DQs, at least one full frame will be serviced. Thus, $B_{ff}(t' + N) = B_{ff}(t') + N$.

Hence, $B_{ff}(t' + N) = B_{ff}(t') + N$

□

Since Lemma 1 shows that ASF-RR with 3DQs is work-conserving for every frame arrive to the 3DQs, Lemma 2 shows that the number of serviced packets is close to the number of packets serviced by an OQ switch.

Lemma 2 $B_{ff}(t') \ge B_D^{OQ}(t') - N^2$

Proof: By induction on:

$t' = \{k + 1 - N, K + 1, \ldots, K + 1 + l \cdot N\}$.

If $t' = k + 1 - N$, $B_{ff}(t') = 0$ and $B_D^{OQ}(t') \le t' \le 1 \le N^2$, so the inequality holds. Now, let the inequality be true for t'.

Case 1: $ff(t') = 0$. Then,

$$B_{ff}(t') = A_{ff}(t') = \sum_{i=1}^{N} N \cdot \left\lceil \frac{A_i(t')}{N} \right\rceil$$

$$\ge \left[\sum_{i=1}^{N} A_i(t') \right] - N(N-1)$$

$$\ge A(t') - N(N-1)$$

$$\ge B_D^{OQ}(t') - N(N-1)$$

$$B_{ff}(t' + N) \ge B_D^{OQ}(t') - N(N-1)$$

$$\ge B_D^{OQ}((t' + N) - N) - N(N-1)$$

$$\ge B_D^{OQ}(t' + N) - N^2$$

Case 2: $ff(t') > 0$. Using Lemma 1,

$$B_{ff}(t' + N) \ge B_{ff}(t') + N$$

$$\ge B_D^{OQ}(t') - N^2 + N$$

$$\ge B_D^{OQ}(t' + N) - N^2$$

□

Since Lemma 2 shows that the number of packet serviced by ASF-RR with 3DQs is closed to the delayed OQ in some sense, Lemma 3 concludes that the queue size for ASF-RR with 3DQs is not bounded by the sum of the queue size for the delayed OQ and a constant. However, for heavy traffic, the queue size for ASF-RR with 3DQs is close to this bound value.

Lemma 3 $q_{ff}(t') - q_D^{OQ}(t') \le N^2 + N - 1$ and

$q(t') - q_D^{OQ}(t') \le N^2$.

Proof: Using the definitions and Lemma 2:

$$q_{ff}(t') - q_D^{oq}(t') = [A_{ff}(t') - B_{ff}(t')] - [A(t') - B_D^{oq}(t')]$$
$$= [A_{ff}(t') - A(t')] + [B_D^{oq}(t') - B_{ff}(t')]$$

Using,

$\max[A_{ff}(t') - A(t')] = N-1$,

and $\max[B_D^{oq}(t') - B_{ff}(t')] = N^2$

Thus, $q_{ff}(t') - q_D^{oq}(t') \le N-1+N^2$
$\le N^2 + N - 1$

Similarly,

$$q(t') - q_D^{oq}(t') = [A(t') - B(t')] - [A(t') - B_D^{oq}(t')]$$
$$= B_D^{oq}(t') - B(t')$$
$$\le B_D^{oq}(t') - B_{ff}(t')$$

Apply lemma 2, thus $q(t') - q_D^{oq}(t') \le N^2$

□

The next two lemmas shows that ASF-RR with 3DQs efficiently uses bandwidth in order to remove the packets from the queues. Since ASF-RR with 3DQs is work-conserving for all frames. Lemma 4 shows that if there are $ff(t')$ full frames at time t', then exactly $ff(t')$ full frames will be served in the next $ff(t')$ cycles. Lemma 5 generalizes this idea.

Lemma 4 If at least one frame arrives to the second stage, then $B_{ff}(t' + q_{ff}(t')) = A_{ff}(t')$.

Proof: If $ff(t') = 0$, then by definition $q_{ff}(t') = 0$ and $B_{ff}(t') = A_{ff}(t')$. Otherwise, $ff(t') \ge 1$ and we can iteratively apply Lemma 1 $ff(t')$ times:

1. $ff(t') \ge 1$, thus $B_{ff}(t'+N) = B_{ff}(t') + N$

(from Lemma 1).

2. $ff(t'+N) \ge ff(t') - 1$, thus
$B_{ff}(t'+2N) = B_{ff}(t') + 2N$

3. $ff(t'+2N) \ge ff(t') - 2$, thus
$B_{ff}(t'+3N) = B_{ff}(t') + 3N$,

etc.

Finally, $ff(t' + [ff(t') - 1] \cdot N) \ge 1$, thus

$B_{ff}(t' + ff(t') \cdot N) = B_{ff}(t') + ff(t') \cdot N = A_{ff}(t')$

.

By definition $q_{ff}(t) = N \cdot ff(t)$ for all t,

Hence $B_{ff}(t' + q_{ff}(t')) = A_{ff}(t')$

□

Lemma 5
$B(t' + q_{ff}(t') + N(2N-2)) \ge A(t')$

Proof: We already know from lemma 4 that $B_{ff}(t' + q_{ff}(t')) = A_{ff}(t')$. Now we need to show that during the next $(2N-2)$ cycles, at least $A(t') - A_{ff}(t')$ cells will be serviced. During these $(2N-2)$ cycles, there are at least $2N-2$ full frames serviced.

Therefore, using lemma 1 and lemma 4:

$$B_{ff}(t' + q_{ff}(t') + N(2N-2)) = A_{ff}(t') + N(2N-2)$$
$$\ge A(t')$$

.

By definition $B_{ff}(t') \le B(t')$

Hence, $B(t' + q_{ff}(t') + N(2N-2)) \ge A(t')$

□

We have shown that $B(t')$ was tracking $A(t')$ with a delay dependent on $q_{ff}(t')$. Since we have linked $q_{ff}(t')$ with $q_D^{oq}(t')$, we can find a first bound on the average delay in the delay for ASF-RR with 3DQs as a function of the average delay in the delayed OQ model.

Theorem 1 The average delay for ASF-RR with 3DQs is less than the average delay for the delayed OQ plus a constant $3N^2 - 2$

Proof: We know that $q_{ff}(t') \le N^2 + N - 1 + q_D^{oq}(t')$ and

$B\left(t' + q_{ff}(t') + N(2N-2)\right) \geq A(t')$ (from Lemma 3 and Lemma5).

Therefore,

$B\left(t' + \left[q_D^{OQ}(t') + N^2 + N - 1\right] + N(2N-2)\right) \geq A(t')$

or

$B\left(t' + q_D^{OQ}(t') + N(3N-1) - 1\right) \geq A(t')$ .But we also know that: $B_D^{OQ}\left(t' + q_D^{OQ}(t')\right) \leq B_D^{OQ}(t') + q_D^{OQ}(t') = A(t')$

.

Hence,

$B\left(t' + q_D^{OQ}(t') + N(3N-1) - 1\right) \geq B_D^{OQ}\left(t' + q_D^{OQ}(t')\right)$

. This implies that the average delay for a cell coming at time t' is the delay that it would have under the delayed OQ algorithm, plus at most $N(3N-1)-1$.

Finally, note that all computations up until now were for a t' that begins a cycle for output k. If we choose any nonnegative integer t, then let $t' = N \cdot \left\lfloor \frac{t-(k+1)}{N} \right\rfloor + k + 1$ be the beginning of the cycle to which t belongs, and let $C = N(3N-1) - 1 + (N-1) = 3N^2 - 2$. We get:

$$\begin{aligned} B\left(t + q_D^{OQ}(t) + C\right) &\geq B\left(t' + \left[A(t) - B_D^{OQ}(t)\right] + C\right) \\ &\geq B\left(t' + A(t') - B_D^{OQ}(t') - (N-1) + C\right) \\ &\geq B\left(t' + q_D^{OQ}(t) + N(3N-1) - 1\right) \\ &\geq A(t') = A(t) \end{aligned}$$

Hence,

$B\left(t + q_D^{OQ}(t) + C\right) \geq B_D^{OQ}\left(t + q_D^{OQ}(t)\right)$ and the result is thus applicable to any time-slot t.

□

Theorem 2 The average delay for ASF-RR with 3DQs is less than the average delay for OQ plus a constant $4N^2 + N - 3$

Proof: We compare the delays for OQ and for the delayed version of OQ. Let $D = N^2 + N - 1$. We'll first show that the delay of any cell in the delayed OQ is less than its delay for OQ plus D.

For any time-slot s, let $s'(s)$ be the time-slot that marks the beginning of the cycle to which s belongs:

$s'(s) = N \cdot \left\lfloor \frac{s-(k+1)}{N} \right\rfloor + k + 1$, and

$s - (N-1) \leq s'(s) \leq s$.

Then, according to the properties of the delayed OQ we have:

$$\begin{aligned} B_D^{OQ}(t) &= \min_{0 \leq s \leq t} \left[A(s) + (t-s)\right] \\ &= \min_{0 \leq s \leq t} \left[A(s'(s)) + (t-s)\right] \\ &\geq \min_{0 \leq s \leq t} \left[A^o\left(s'(s) - N^2\right) + (t-s)\right] \\ &\geq \min_{0 \leq s \leq t} \left[A^o(s-D) + (t-s)\right] \end{aligned}$$

Let $\tau = s - D$. Since packets don't arrive before time-slot 0:

$$\begin{aligned} B_D^{OQ}(t) &\geq \min_{-D \leq \tau \leq t-D} \left[A^o(\tau) + (t-D) - \tau\right] \\ &\geq \min_{0 \leq \tau \leq t-D} \left[A^o(\tau) + (t-D) - \tau\right] \\ &\geq B^{OQ}(t-D) \end{aligned}$$

Since the delayed OQ and OQ are both FCFS, the difference of delay for each cell between those two systems will be at most D. Hence, the difference of average delay between ASF-RR with 3DQs and OQ will be at most $3N^2 - 2 + N^2 + N - 1 = 4N^2 + N - 3$.

□

It is possible to improve this bound using a difference algorithm that would take into account the number of cells present in the non-full frames. However, this would increase the complexity and the communication in the switch, and we believe that the trade-off is not worth it.

C. ASF-RR with 3DQs has the same throughput as OQ

Let's first provide a few definitions. Consider a switch with traffic arrival rates (ρ_{ik}) (from EI i to EO k), and total queueing size $Q(t)$, where t is the current time-slot

1. The load of the arrival traffic is:

$\rho = \max\left(\max_k\left(\sum_{i=1}^{N}\rho_{ik}\right), \max_i\left(\sum_{k=1}^{N}\rho_{ik}\right)\right)$. The arrival traffic is said to be admissible if $\rho < 1$.

2. The switch is said to be strongly stable if $\limsup_{t\to\infty}\left(E\left[Q(t)\right]\right) < \infty$ [17].

3. The switch is said to have 100% throughput if it is strongly stable whenever the arrival traffic is admissible.

We have seen that there exists a bound average delay difference between ASF-RR with 3DQs and OQ switch. As a sequence, ASF-RR with 3DQs has the same throughput as OQ, as Theorem 3 illustrates.

Theorem 3 ASF-RR with 3DQs and OQ have the same throughput.

Proof: Let $Q^{OQ}(t)$ and $Q^{ASF}(t)$ be the total queueing size at time-slot t in a switch implementing OQ and ASF-RR with 3DQs, respectively. Also, let $C = 4N^3 + N^2 - 2N$, and assume that the cycle for EO k begin at $t'(t)$, with $t-(N-1) \le t'(t) \le t$. Using Lemma 3 and the proof for Theorem 2, we get:

$$\begin{aligned} q(t) &\le q(t'(t)) + N(N-1) \\ &\le q_D^{OQ}(t'(t)) + N^2 + N(N-1) \\ &\le \left[A(t'(t)) - B_D^{OQ}(t'(t))\right] + 2N^2 - N \\ &\le \left[A(t) - B_D^{OQ}(t)\right] + (N-1) + 2N^2 - N \\ &\le A(t) - B^{\infty}(t-D) + 2N^2 - 1 \\ &\le A^{\circ}(t) - B^{\infty}(t) + D + 2N^2 - 1 \\ &\le q^{OQ}(t) + 3N^2 + N - 2 \end{aligned}$$

Taking into account both the buffering in the 3DQs for the N external outputs and the buffering in the VOQ1s from the N external inputs, we get (using $q_k(t) \equiv q(t)$ for EO k):

$$\begin{aligned} Q^{ASF}(t) &\le N\sum_{k=1}^{N} q_k^{ASF}(t) + N \cdot N^2 \\ &\le N\sum_{k=1}^{N} q_k^{OQ}(t) + N \cdot \left(3N^2 + N - 2\right) + N^3 \\ &\le Q^{OQ}(t) + C \end{aligned}$$

Thus, $Q^{OQ}(t) \le Q^{ASF}(t) \le Q^{OQ}(t) + C$, since OQ is work-conserving. Therefore, they have the same throughput.

□

Property 4 Assume that the arrival traffic patterns from EI i to EO k are (σ, ρ_{ik})-upper constrained. If the traffic is admissible, then ASF-RR with 3DQs has 100% throughput.

Proof: Let $A_{ik}(n)$ be the cumulative number of arrivals for cells going from i to k at time-slot n, and let $A_k(n) = \sum_{i=1}^{N} A_{ik}(n)$. Then by definition, for all $1 \le i,k \le N$ and $0 \le m \le n$, we have

$A_{ik}(n) - A_{ik}(m) \le \rho_{ik} \cdot (n-m) + \sigma$. This implies that $A_k(n) - A_k(m) \le \rho \cdot (n-m) + N \cdot \sigma$, thus arrivals to any external output are $(N\sigma, \rho)$-upper constrained. Since $\rho < 1$, the maximum queue length size in OQ is $N\sigma$ [11]. The result follows using Theorem 3

□

V. CONCLUSION

Over the last few years, there have been many results that show the conditions under which single-stage crossbar switch with input queues and no speedup can achieve 100% throughput.

The two stage-switch introduced by C.S. Chang *et al.* [10] achieves a 100% throughput as well as a bound on the delay between it and an output queued switch. This is achieved without speedup and without a complicated scheduling algorithm, and therefore represents an important step towards efficient, high capacity switches with delay guarantees. In its simplest form the two-stage switch miss-sequence packets. To maintain the sequence of packets in the flow and bound the miss-sequence packets, Keslassy *et al.* [11] proposed the Full Frame First algorithm that prevent miss-sequencing while maintaining the throughput and delay properties of the basic switch. However, it clearly introduces more complexity and if some packets loss, then this scheme will accumulate the miss-sequence packets in the buffers

In order to solve these problems, we proposed the scheme called Adaptation Size Frame-Round Robin with 3DQs mechanism (ASF-RR with 3DQs). This scheme is simple form and can achieve 100% throughput guarantees and bound the delay time as well. This scheme has the following advantages:

(i) The scheme can overcome head-of-line blocking and enable high throughput and provide delay

guarantees for a broad class of arrival processes.

(ii) It will not accumulate the miss-sequence packets in the buffers if some packets loss.

(iii) The average packet delay for ASF-RR with 3DQs is bounded by the sum of the average packet delay for OQ (with the same traffic) and a constant delay. No delay bounds exist for *i*SLIP, which is widely used practical heuristic for single-stage crossbar switches.

(iv) ASF-RR with 3DQs does not require much information to be sent between each internal input and external output.

(v) ASF-RR with 3DQs seems well suited to optical switch fabrics base on technologies such as MEMS, VCSE, tunable lasers, etc.

However, ASF-RR with 3DQs has some drawbacks.

(i) ASF-RR with 3DQs uses two switching stages instead of one. This is similar to using a crossbar switch with speedup of two that normally limit the speed of the system- *i.e.* the bandwidth limit of the memories at each stage, and the scheduler that shall be run at the same speed as the external line.

(ii) ASF-RR with 3DQs needs double amount of buffers (first-and-second stage) instead of one buffer system.

(iii) ASF-RR with 3DQs uses N^3 3DQs in the internal buffer, while single stage switch usually use only N^2 VOQs, thus requiring more pointers, and a more complicated buffer management algorithm.

We believe that the most interesting application of the two-stage switch is for use as the optical switching fabric in an otherwise electronic Internet router. Its power consumption, its size and the need for a complex scheduler generally limit the switch fabric in a router while optics can reduce both size and power.

REFERENCES

[1] M. Ajmone Marsan, A. Bianco, P. Giaccone, E. Leonardi and F. Neri, "Packet scheduling in input-queued cell-based switches," *IEEE Infocom'01*, Anchorage, Alaska, Apr. 2001.

[2] N. McKeown, "iSLIP: A scheduling algorithm for input-queued switches," *IEEE Trans. on Networking*, Vol. 7, No.2, Apr. 1999.

[3] T. E. Anderson, S. S. Owicki, J. B. Saxe, and J. Walrand, "High speed switch scheduling for local area networks," *ACM Trans. On Computer Systems*, Vol. 11, No.4, pp.319-352, Nov. 1993.

[4] E. Leonardi, M. Mellia, F. Neri and M. A. Marsan, "On the stability of input-queued switches with speed-up," *IEEE/ACM Trans. on Networking*, Vol. 9, No. 1, pp. 104-118, Feb. 2001.

[5] J. G. Dai and B. Praphakar, "The throughput of data switches with and without speedup," *IEEE Infocom'00*, Vol. 2, pp. 556-564, Tel Aviv, Israel, Mar. 2000.

[6] N. McKeown, A. Mekkittikul, V. Anantharam and J. Walrand, "Achieving 100% throughput in an input-queued switch," *IEEE Trans. on Communications*, Vol. 47, No. 8, Aug. 1999.

[7] I. Keslassy, S. T. Chuang, and N. McKeown, "A load-balanced switch with an arbitrary number of linecards," *IEEE Infocom'04*, Hong Kong, Mar. 2004.

[8] C. S. Chang, D. S. Lee and Y. J. Shih, "Mailbox switch: a scalable two-stage switch architecture for conflict resolution of ordered packets," *IEEE Infocom'04*, Hong Kong, Mar. 2004.

[9] I. Keslassy, S. T. Chuang, K. Yu, D. Miller, M. Horowitz, O. Solgaard and N. McKeown, "Scaling internet routers using optics," *ACM SIGCOM'03*, Karlsruhe, Germany, Aug. 2003.

[10] C. S. Chang, D. S. Lee and C. Y. Yue,

"Providing guaranteed rate services in the load balanced Birkhoff-von Neumann switches," *IEEE Infocom'03*, 2003

[11] I. Keslassy and N. McKeown, "Maintaining packet order in two-stage switches," *IEEE Infocom'02*, Jun. 2002.

[12] L. G. Valiant, "A scheme for fast parallel communication," *SIAM Journal on Computing*, Vol. 11, No. 2, pp. 350-361, 1982

[13] E. Blanton and M. Allman, "On making TCP more robust to packet reordering," *ACM Computer Communication Review*, Vol. 32, No. 1, pp. 20-30, Jan. 2002.

[14] M. Fomenkov, K. Keys, D. Moore and K. Claffy, "A longitudinal study of internet traffic from 1998-2001: a view from 20 high performance sites," *Proc. Of WISICT'04*, Cancun, Mexico, Jan. 2004.

[15] C. S. Chang, D. S. Lee and C. M. Lien, "Load balanced Birkhoff-von Neumann switches, Part II: multi-stage buffering," *Computer Communications*, Vol. 25, No. 6, pp. 623-634, 2002.

[16] I. Keslassy, C. S. Chang, N. McKeown and D. S. Lee, "Optimal load-balancing," *IEEE Infocom'05*, Mar. 2005.

ประสิทธิภาพการป้องกันการจัดเรียงตัวใหม่ของแพ็กเก็ตในเราเตอร์แบบถ่วงสมดุลย์

The Performance on the Prevention of Packet Reordering in Load-Balanced Router

กนกวรรณ ไวยนนท์ (Kanokwan Waiyanon)*

พิเชษฐ เชี่ยวธนะกุล (Bhichate Chiewthanakul)**

บทคัดย่อ

ในงานวิจัยนี้อยู่ภายใต้การสร้างแบบสวิทช์สองขั้นตอน (two-stage switch) ซึ่งงานทดลองนี้ได้นำเสนอขั้นตอนวิธีที่เรียกว่า ASF-RR ร่วมกับคิวสามมิติ (Adaptation-Size-Frame-RR with 3 Dimension Queues) ที่สามารถป้องกันปัญหาการจัดเรียงตัวที่ผิดลำดับของแพ็กเก็ตในเราเตอร์แบบถ่วงสมดุลย์ อันเนื่องมาจากแพ็กเก็ตต่างๆของสายข้อมูลเดียวกันอาจถูกลำเลียงในเส้นทางที่ต่างกัน ในงานวิจัยนี้ประสิทธิภาพได้ถูกศึกษาในเทอมของค่าหน่วงของการป้องกันการจัดเรียงตัวที่ผิดลำดับของแพ็กเก็ตในเราเตอร์แบบถ่วงสมดุลย์ โดยใช้โปรแกรม MATLAB เพื่อจำลองและเปรียบเทียบค่าหน่วงที่เกิดขึ้นระหว่าง ASF-RR ร่วมกับคิวสามมิติ กับ FFF (Full Frame First) ผลการจำลองแสดงให้เห็นว่าค่าหน่วงเฉลี่ยระหว่าง ASF-RR ร่วมกับคิวสามมิติ และ FFF มีความแตกต่างกันเพียงเล็กน้อย แต่ ASF-RR ร่วมกับคิวสามมิติ จะไม่มีการสะสมแพ็กเก็ตที่ผิดลำดับในบัฟเฟอร์ถ้าหากมีบางแพ็กเก็ตสูญหายเหมือนใน FFF

ABSTRACT

In this paper, building on the two-stage switch, an algorithm called ASF-RR with 3DQs (Adaptation-Size-Frame-RR with Three Dimensional Queues) is proposed.This algorithm can prevent the problem of packets out of sequence in the load-balanced router in which different packets of the same flow may be taken into different paths. In this paper, the performance is studied in the term of delay of the prevention of packets reordering in load-balanced router. By using MATLAB to simulate and to compare the delay between ASF-RR with 3DQs and FFF (Full Frame First).The simulations show that the average delays between ASF-RR with 3DQs and FFF are slightly difference but ASF-RR with 3 DQs will not accumulate the miss-sequence packets in the buffers if some packets loss like in FFF.

คำสำคัญ : เราเตอร์แบบถ่วงสมดุลย์ คิวเอาท์พุทเสมือน การจัดเรียงตัวใหม่ของแพ็กเก็ต

Key Words: Load-Balanced Router, Virtual Output Queues, Packet Reordering

---

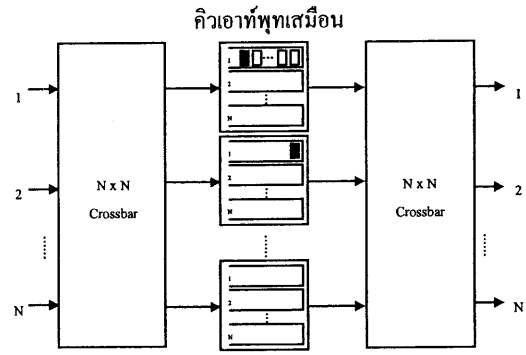
*นักศึกษาระดับบัณฑิตศึกษา หลักสูตรวิศวกรรมศาสตร์มหาบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น

**ผู้ช่วยศาสตราจารย์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น

บทนำ

เราเตอร์หลักในอินเทอร์เน็ตปัจจุบันได้นำคิวที่ใช้งานร่วมกันระหว่างคิวอินพุทและคิวเอาท์พุท (combined input and output queueing) มาใช้งานร่วมกับตัวจัดกำหนดการส่วนกลาง (centralized scheduler) มีขั้นตอนวิธีของตัวจัดกำหนดการส่วนกลางหลายวิธี (Marsan et al,2001; McKeown, 1999; Anderson, 1993) อย่างไรก็ตามแม้ว่าขั้นตอนวิธีเหล่านี้โดยทฤษฎีสามารถที่จะรับประกันค่าประสิทธิผลได้ 50-100% (Leonardi et al.,2001; Dai and Praphakar, 2000; McKeown, 1999) แต่ในทางปฏิบัติไม่สามารถเป็นไปได้เมื่ออัตราเร็วของสายสัญญาณและจำนวนพอร์ตเพิ่มมากขึ้นเนื่องจากความซับซ้อนของขั้นตอนวิธีและ/หรือการเพิ่มอัตราเร็วของบัฟเฟอร์ของหน่วยความจำ

ไม่นานมานี้ได้มีผู้นำเสนอวิธีใหม่คือการไม่ใช้การจัดกำหนดการ (scheduling) โดยการใช้สถาปัตยกรรมของสวิทช์แบบถ่วงสมดุลย์ (Keslassy et al., 2001; Chang et al., 2004; Keslassy et al., 2003; Chang et al., 2003; Keslassy and McKeown, 2002) สถาปัตยกรรมนี้อาศัยหลักการถ่วงสมดุลย์แพ็กเก็ตอย่างสม่ำเสมอภายในเราเตอร์ก่อนที่จะส่งต่อไปยังปลายทางที่ถูกต้อง แนวคิดนี้ครั้งแรกได้มาจาก Valiant (1982) ส่วน Chang et al. (2003) ได้แสดงให้เห็นว่าเราเตอร์แบบถ่วงสมดุลย์โดยใช้คิวเอาท์พุทเสมือนไม่ต้องการตัวจัดกำหนดการและสามารถรับประกันค่าประสิทธิผลได้ 100% เราเตอร์แบบถ่วงสมดุลย์จึงไม่เกิดปัญหาหลักสองอย่างที่เกิดขึ้นกับเราเตอร์ในปัจจุบัน ได้แก่ปัญหาของตัวจัดกำหนดการส่วนกลางและปัญหาการรับประกันค่าประสิทธิผล ซึ่งทำให้สถาปัตยกรรมของเราเตอร์แบบถ่วงสมดุลย์นี้เป็นที่น่าสนใจต่อการศึกษาอย่างมาก อย่างไรก็ตามเราเตอร์แบบถ่วงสมดุลย์มีปัญหาหลายอย่างที่สำคัญ เช่น การจัดเรียงตัวใหม่ของแพ็กเก็ตและความต้องการจัดค่าคอนฟิกูเรชั่นใหม่ของสวิทช์แฟบริคบ่อยครั้ง ผลที่สุดก็คือทีซีพี (TCP: โพรโทคอลควบคุมการส่งผ่าน) ทำงานได้ไม่ดีเมื่อแพ็กเก็ตที่เรียงตัวผิดลำดับมาถึงปลายทาง แพ็กเก็ตที่เรียงตัวผิดลำดับสามารถเป็นตัวบ่งชี้ถึงความเสียหาย และทำให้เกิดการส่งใหม่ที่ไม่จำเป็นและทำให้เกิดการหมดเวลาของทีซีพี (TCP timeouts) (Blanton and Allman, 2002) การส่งใหม่และการหมดเวลาของทีซีพีเป็นสาเหตุให้ค่าประสิทธิผลของทีซีพีลดลงและเพิ่มค่าหน่วงของแพ็กเก็ต เพราะฉะนั้นเนื่องจากการจราจร (traffic) ของทีซีพีเป็นการจราจรส่วนใหญ่ของอินเทอร์เน็ต ดังนั้นเราเตอร์จะต้องไม่จัดเรียงแพ็กเก็ตใหม่ภายในสายข้อมูลเดียวกัน (Fomenkov et al., 2004)



ภาพที่ 1 ตัวอย่างการจัดเรียงตัวใหม่ของแพ็กเก็ตในเราเตอร์แบบถ่วงสมดุลย์

ภาพที่ 1 แสดงถึงความเป็นไปได้ของการจัดเรียงตัวใหม่ของแพ็กเก็ตในเราเตอร์แบบถ่วงสมดุลย์ ในตัวอย่างนี้แพ็กเก็ตสีขาวแสดงถึงแพ็กเก็ตที่เข้ามาก่อนแพ็กเก็ตสีดำ แพ็กเก็ตทั้งหมดที่นำมาพิจารณามีปลายทางไปที่เอาท์พุทที่หนึ่งโดยนิยามในช่วง N สล็อทเวลา (time slots) แรก แต่ละอินพุท

จะเชื่อมต่อกับคิวเอาท์พุทเสมือนที่หนึ่งของไลน์คาร์ดที่หนึ่ง หนึ่งครั้ง สมมติให้แต่ละอินพุทยกเว้นอินพุทที่หนึ่ง รับแพ็กเก็ตสีขาวเข้ามาและส่งแพ็กเก็ตนั้นไปยังคิวเอาท์พุทเสมือนที่หนึ่งของไลน์คาร์ดที่หนึ่ง หลังจาก N สล็อทเวลา คิวเอาท์พุทเสมือนที่หนึ่งจะได้รับแพ็กเก็ตสีขาวจำนวน N-1 แพ็กเก็ตจากอินพุท 2,...,N และมีเพียงแพ็กเก็ตเดียวเท่านั้นที่จะได้รับบริการ ดังนั้นจึงมีแพ็กเก็ตสีขาวจำนวน N-2 แพ็กเก็ตที่คงเหลือในคิวเอาท์พุทเสมือนที่หนึ่ง ต่อมาอินพุทที่หนึ่งรับแพ็กเก็ตสีดำเข้ามาจำนวนสองแพ็กเก็ตติดต่อกันเมื่ออินพุทที่หนึ่งได้รับการเชื่อมต่อแล้วจะส่งแพ็กเก็ตสีดำทั้งสองนี้ไปยังคิวเอาท์พุทเสมือนโดยที่แพ็กเก็ตสีดำลำดับที่หนึ่งถูกส่งไปยังคิวเอาท์พุทเสมือนที่หนึ่งของไลน์คาร์ดที่หนึ่ง และแพ็กเก็ตสีดำลำดับที่สองถูกส่งไปยังคิวเอาท์พุทเสมือนที่หนึ่งของไลน์คาร์ดที่สอง แพ็กเก็ตสีดำลำดับที่สองนั้นอยู่ในคิวเพียงแพ็กเก็ตเดียว ขณะที่แพ็กเก็ตสีดำลำดับที่หนึ่งมีแพ็กเก็ตสีขาวจำนวน N-2 แพ็กเก็ตอยู่ข้างหน้า เพราะฉะนั้นแพ็กเก็ตสีดำลำดับที่สองจะได้รับบริการก่อน นั่นคือแพ็กเก็ตสีดำลำดับที่สอง จะมาถึงเอาท์พุทที่หนึ่งก่อนและออกจากเราเตอร์ไปก่อนแพ็กเก็ตสีดำลำดับที่หนึ่ง ผลที่สุดก็คือทั้งสองแพ็กเก็ตมีการจัดเรียงตัวกันใหม่

มีสองวิธีหลักสำหรับการป้องกันการจัดเรียงตัวใหม่ของแพ็กเก็ต วิธีแรกคือการจำกัดขอบเขตแพ็กเก็ตที่ผิดลำดับและใช้บัฟเฟอร์ที่จำกัดจำนวนที่เอาท์พุทจัดเรียงแพ็กเก็ตให้ถูกต้องตามลำดับ วิธีที่สองคือการป้องกันไม่ให้แพ็กเก็ตเรียงผิดลำดับตลอดเส้นทางภายในเราเตอร์ ดังได้เห็นจากตัวอย่างข้างต้นแล้วว่าการจัดเรียงตัวใหม่เกิดขึ้นเมื่อคิวเอาท์พุทเสมือนมีความยาวแตกต่างกันและจำนวนของการจัดเรียงตัวจะเพิ่มขึ้นเมื่อความยาวของคิวเอาท์พุทเสมือนเพิ่มขึ้น เพราะฉะนั้นขั้นตอนวิธีส่วนใหญ่ซึ่งป้องกันการจัดเรียงตัวใหม่จะพยายามจำกัดขอบเขตหรือป้องกันไม่ให้ความยาวของคิวเอาท์พุทเสมือนแตกต่างกัน

Chang et al.(2002) ได้เสนอขั้นตอนวิธีในการแก้ปัญหาแพ็กเก็ตเรียงตัวผิดลำดับสองขั้นตอนวิธี โดยใช้หลักการของการจำกัดจำนวนแพ็กเก็ตที่ผิดลำดับ ทั้งสองขั้นตอนวิธีนั้นมีใช้ในเราเตอร์ที่ส่งแพ็กเก็ตแบบขนานและใช้บัฟเฟอร์ขนาดเล็กที่อินพุท ขั้นตอนวิธีแรกเรียกว่า FCFS (First Come First Serve) ใช้บัฟเฟอร์ควบคุมความต่อเนื่องของข้อมูลในแต่ละคิวเอาท์พุทเสมือนเพื่อให้แน่ใจว่าได้ลำดับที่ถูกต้องในการออกไปจากคิวเอาท์พุทเสมือน ขั้นตอนวิธีที่สองเรียกว่า EDF(Earliest Deadline First) เป็นการจัดแพ็กเก็ตให้สอดคล้องกับเวลาที่จะออกไปในเราเตอร์อุดมคติ อย่างไรก็ตามทั้งสองขั้นตอนวิธีเป็นไปได้ยากในทางปฏิบัติ กลไกการควบคุมความต่อเนื่องของข้อมูลใน FCFS อาจต้องการจำนวน N ครั้งในการเขียนหน่วยความจำต่อสล็อทเวลา และ EDF ต้องการที่จะคืนแพ็กเก็ตที่มีค่าเวลาที่น้อยสุดจากคิว ทำให้เป็นการยากที่จะนำไปใช้ในเราเตอร์ที่มีประสิทธิภาพสูงได้ วิธีที่สองในการป้องกันการจัดเรียงตัวใหม่โดยทำให้แพ็กเก็ตเรียงลำดับถูกต้องตลอดเส้นทางในเราเตอร์ถูกนำมาใช้โดย Chang et al.(2003) Keslassy and McKeown (2002) และ Keslassy et al (2005)

Chang et al. (2003) ได้เสนอขั้นตอนวิธีซึ่งจำกัดขนาดบัฟเฟอร์ในคิวเอาท์พุทเสมือน แพ็กเก็ตได้รับการประกันในการออกจากเราเตอร์ตามลำดับที่ถูกต้อง สมมติว่าในแต่ละช่วงเวลาเฟรมของขั้นตอนวิธีนี้ การมาถึงของแพ็กเก็ตที่มีปลายทางไปยังเอาท์พุทที่กำหนดจะถูกควบคุม อย่างไรก็ตามคุณสมบัตินี้ไม่เป็นที่ต้องการ

ในอินเทอร์เน็ตทั่วไป

Keslassy and McKeown (2002) ได้เสนอขั้นตอนวิธีที่ใช้บัฟเฟอร์ในขั้นตอนอินพุทแล้วส่งแพ็กเก็ตไปเก็บในคิวเอาท์พุทเสมือนโดยจำแนกตามอินพุท คิวเอาท์พุทเสมือน และเอาท์พุท การใช้คิวเช่นนี้ทำให้ขั้นตอนวิธีสามารถรับประกันว่าแพ็กเก็ตมาถึงเอาท์พุทตามลำดับที่ถูกต้อง อย่างไรก็ตามขั้นตอนวิธีนี้ต้องการระบบการจัดการเกี่ยวกับคิวที่ซับซ้อนมากกว่าและต้องการการติดต่อสื่อสารของข้อมูลสถานะระหว่างคิวเอาท์พุทเสมือนและเอาท์พุท

Keslassy et al.(2005) ได้เสนอขั้นตอนวิธีที่เรียกว่า FOFF (Full Ordered Frames First) ยอมให้มีการจัดเรียงตัวใหม่ของแพ็กเก็ตภายในเราเตอร์ได้จำนวนหนึ่งและก่อนออกจากเราเตอร์จะจัดเรียงแพ็กเก็ตที่ผิดลำดับนี้ให้ถูกต้องโดยใช้บัฟเฟอร์สำหรับจัดเรียงลำดับที่เอาท์พุท (resequencing buffer) เนื่องจากกลไกการควบคุมนี้ยอมให้คิวที่ไม่ว่างสามารถส่งแพ็กเก็ตได้แม้ไม่มีคิวใดที่มีแพ็กเก็ตเต็มเฟรมก็ตาม ดังนั้น FOFF จึงหลีกเลี่ยงการที่คิวนั้นจะไม่ได้รับการบริการ (starvation) แต่ยอมให้คิวเอาท์พุทเสมือนมีความยาวที่แตกต่างกันได้ ทำให้เพิ่มจำนวนแพ็กเก็ตที่ผิดลำดับมากขึ้นในคิวเอาท์พุทเสมือน

เนื่องจากขั้นตอนวิธีเหล่านี้นำเสนอปัญหาที่สำคัญซึ่งไม่เป็นสิ่งที่ต้องการในอินเทอร์เน็ต ดังนั้นวัตถุประสงค์ของงานวิจัยนี้คือหาวิธีที่แตกต่างเพื่อที่จะสามารถรับประกันการที่แพ็กเก็ตจะออกจากเราเตอร์อย่างถูกต้องตามลำดับและเพื่อป้องกันแพ็กเก็ตจัดเรียงตัวผิดลำดับ วิธีดังกล่าวนี้สามารถรับประกันค่าประสิทธิผลและค่าหน่วงสำหรับทุกรูปแบบการจราจร

สถาปัตยกรรมของสวิทช์ (Switch Architecture)

สมมติฐาน

ในงานวิจัยนี้ สมมติให้ทุกแพ็กเก็ตที่เข้ามาที่มีความยาวแตกต่างกันจะถูกแบ่งออกให้มีขนาดคงที่หรือเรียกว่าแพ็กเก็ต และจะถูกนำมารวมตัวกันอีกครั้งเมื่อออกจากเราเตอร์ เส้นทางเชื่อมต่อจะเชื่อมต่อไปยังเราเตอร์ภายในไลน์คาร์ดมีจำนวนเท่ากับ N ซึ่งเท่ากับไลน์คาร์ดอินพุทและเอาท์พุท และสมมติให้ทุกไลน์คาร์ดมีอัตราเร็วของการส่งข้อมูลเท่ากันและเวลาจะถูกแบ่งเป็นสล็อท ดังนั้นจะมีอย่างมากเพียงหนึ่งแพ็กเก็ตเท่านั้นที่เดินทางเข้ามายังพอร์ตอินพุทและอย่างมากเพียงหนึ่งแพ็กเก็ตเท่านั้นที่เดินทางออกไปยังพอร์ตเอาท์พุทในแต่ละสล็อทเวลาและสมมติว่าในระยะเริ่มต้นยังไม่มีแพ็กเก็ตใดๆอยู่ในเราเตอร์

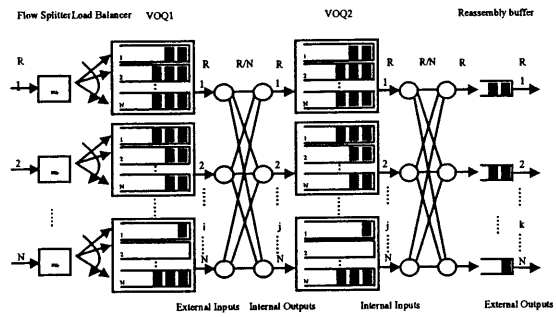
สถาปัตยกรรมของเราเตอร์แบบถ่วงสมดุลย์

สถาปัตยกรรมของสวิทช์ที่นำเสนอได้พัฒนามาจาก Keslassy and McKeown (2002) และแสดงไว้ในภาพที่ 2 สถาปัตยกรรมของเราเตอร์แบบถ่วงสมดุลย์ทำงานบนสวิทช์ที่มีอัตราเร็วเท่ากันสองตัวและบัฟเฟอร์สามชุด บัฟเฟอร์ชุดแรกช่วยจำกัดขอบเขตจำนวนแพ็กเก็ตที่เรียงตัวผิดลำดับ สวิทช์ตัวแรกใช้เพื่อป้องกันการขัดขวางแพ็กเก็ตหัวแถว (head of line blocking) โดยกระจายแพ็กเก็ตในรูปแบบวนรอบ (round-robin) ไปยังบัฟเฟอร์ชุดที่สองหรือคิวเอาท์พุทเสมือนที่สอง ดังนั้นสวิทช์ตัวแรกจะถ่วงสมดุลย์แก่การจราจรที่เข้ามา แพ็กเก็ตจะถูกส่งข้ามสวิทช์ตัวที่สอง ไปยังบัฟเฟอร์ชุดที่สาม ซึ่งเป็นชุดสุดท้ายใช้สำหรับการรวมตัวกลับคืนของแพ็กเก็ตก่อนจะออกจากเราเตอร์

สวิทช์แต่ละตัวมีการจัดตั้งค่าให้มีการวนรอบที่เหมือนกัน นั่นคือ ณ เวลา t อินพุท i ของสวิทช์แต่ละตัวได้ต่อกับเอาท์พุท $[(i+t-1) \bmod N]+1$ ดังนั้นสวิทช์แต่ละตัวได้ต่อเชื่อมอินพุทใด ๆ เข้ากับเอาท์พุทใด ๆที่

$1/N$ –th ของเวลาพอดี โดยไม่สนใจการจราจรที่เข้ามา สวิทช์ทั้งสองเป็นสวิทช์แฟบริคแบบ NxN ที่มีอัตราเร็วคงที่และเท่ากันและมีการจัดรูปแบบการเชื่อมต่อใหม่ทุกสล็อทเวลา อย่างไรก็ตามถ้าจะนำมาใช้กับ สวิทช์แฟบริคแสง ข้อจำกัดนี้จะเป็นอุปสรรคต่อความสามารถในการเพิ่มความเร็วสูง ระบบที่ไม่มีการจัดรูปแบบใหม่เป็นระบบที่ง่ายและมีความน่าเชื่อถือมากกว่าระบบที่มีการจัดรูปแบบใหม่บ่อยครั้ง ดังนั้น สวิทช์แต่ละตัวจะถูกแทนที่ด้วยเมชคงที่ (fixed mesh) ของช่องสัญญาณแสง (optical channels)

ข้อสังเกตอันแรก คือสามารถแทนที่เมชทั้งสองขั้นตอน ด้วยช่องสัญญาณคงที่ N^2 ช่องสัญญาณ ที่อัตราเร็ว R/N ดังแสดงในภาพที่ 2 ดังนั้นเมชนี้ได้นำมาใช้แทนที่สวิทช์ อัตราเร็วที่ได้ระหว่างสวิทช์ยังคงเหมือนเดิม ดังนั้นสวิทช์ยังคงเป็นสวิทช์ที่มีอัตราเร็วคงที่และเท่ากัน ข้อสังเกตที่สองพบว่าสามารถแทนที่เมชสองขั้นตอนด้วยเมชเพียงขั้นตอนเดียวที่มีอัตราเร็วเพิ่มขึ้นเป็น 2 เท่าดังแสดงในภาพที่ 3ก เนื่องจากทุกแพ็กเก็ตเดินทางข้ามเมชสองครั้ง แต่ละครั้งด้วยอัตราเร็ว R/N ดังนั้นอัตราเร็วของช่องสัญญาณทั้งหมดจะเป็น $2R/N$ หลังจากที่แพ็กเก็ตหนึ่งๆข้ามสวิทช์เป็นครั้งแรกแล้วจะถูกเก็บไว้ในไลน์คาร์ดคิวเอาท์พุทเสมือนจากนั้นแพ็กเก็ตจะข้ามสวิทช์อีกครั้งเพื่อออกไป ยังไลน์คาร์ดเอาท์พุท ข้อสังเกตที่สาม เมชที่มีช่องสัญญาณแสงสามารถถูกแทนที่ได้ด้วย AWGR (Arrayed Wave guide Grating Router) ดังแสดงในภาพที่3ข อินพุท i ส่งออกไป N ช่องสัญญาณบนไฟเบอร์ขาออก แต่ละช่องสัญญาณที่แตกต่างกัน λ_j ได้ถูกส่งออกด้วยอัตราเร็ว $2R/N$ บนความยาวคลื่นที่เฉพาะเจาะจง λ_w AWGR เป็นอุปกรณ์



ภาพที่ 2 สถาปัตยกรรมของเราเตอร์แบบถ่วงสมดุลย์โดยใช้เมชสองขั้นตอน

สลับช่องสัญญาณทางแสงแบบพาสสีฟ แต่ละช่องสัญญาณของอินพุทที่กำหนดให้จะออกไปที่เอาท์พุทต่าง ๆ กัน ผลที่ได้คือระบบมีพฤติกรรมดังเช่นเมชเดี่ยว (single mesh) และข้อดีที่สำคัญคือการลดจำนวนของไฟเบอร์ลง จาก N^2 เหลือเพียง $2N$

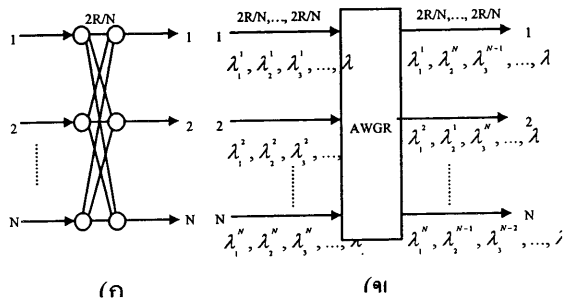
คำจำกัดความ

สถาปัตยกรรมของสวิทช์ประกอบด้วยสวิทช์สองขั้นตอน อินพุทของขั้นตอนที่หนึ่งเรียกว่าอินพุทภายนอก (external inputs:EIs) กำหนดให้เป็น $i = 1,...,N$ เอาท์พุทของขั้นตอนที่หนึ่งเรียกว่าเอาท์พุทภายใน (internal outputs:IOs) ซึ่งจะเป็นอินพุทให้แก่สวิทช์ขั้นตอนที่สอง เรียกว่าอินพุทภายใน (internal inputs:IIs) เอาท์พุทภายในและอินพุทภายในจะใช้แทนซึ่งกันและกันในงานวิจัยนี้ และกำหนดให้เป็น $j = 1,...,N$ และสุดท้ายเอาท์พุทของขั้นตอนที่สองหรือเรียกว่าเอาท์พุทภายนอก (external outputs:EOs) กำหนดให้เป็น $k = 1,...,N$

เส้นทางของแพ็กเก็ตภายในสวิทช์

1. สายข้อมูลจะถูกแบ่งออกโดยตัวแยก (flow splitter) และลาเบลแต่ละแพ็กเก็ตในอินพุทภายนอก i

ซึ่งเป็นของสายข้อมูล F_{ik} เมื่อ k เป็นเอาท์พุทภายนอก ซึ่งเป็นปลายทางที่แพ็กเก็ตจะออกไป ดังนั้นจึงมีเอาท์พุทภายนอก ที่แตกต่างกันเป็นจำนวน N สายข้อมูลที่เป็นไปได้ต่อหนึ่งอินพุทภายนอก



ภาพที่ 3 สถาปัตยกรรมของเราเตอร์แบบถ่วงสมดุลย์ โดยใช้ (ก) เมชเดี่ยว (ข) AWGR

2. ตัวถ่วงสมดุลย์ส่งแพ็กเก็ตทั้งหมดจาก F_{ik} ไปยัง คิวเอาท์พุทเสมือนที่หนึ่ง N ตัว แบบวนรอบ (round-robin) ตัวอย่างเช่นแพ็กเก็ตแรกจากสายข้อมูลที่กำหนดจะถูกส่งออกไปที่คิวเอาท์พุทเสมือนที่หนึ่งสำหรับเอาท์พุทภายในที่หนึ่ง แพ็กเก็ตที่สองถูกส่งออกไปที่คิวเอาท์พุทเสมือนที่หนึ่งสำหรับเอาท์พุทภายในที่สอง และต่อๆไป อย่างเป็นอิสระจากเวลาที่เข้ามา ถ้าสายข้อมูลมีแพ็กเก็ตไม่ครบจำนวน N จะมีการเติมแพ็กเก็ตว่างที่มีขนาดเดียวกันเข้าไปที่คิวเอาท์พุทเสมือนที่หนึ่ง จนครบจำนวน N เพราะว่าตัวถ่วงสมดุลย์ไม่จำเป็นต้องเข้าจังหวะ (synchronize) กับลำดับการทำงานของสวิทช์ขั้นตอนที่หนึ่ง แพ็กเก็ตที่มาถึงจะถูกเก็บไว้ในบัฟเฟอร์และไม่จำเป็นที่จะต้องออกจากคิวเอาท์พุทเสมือนที่หนึ่งในทันที สังเกตว่าความยาวของคิวในแต่ละคิวเอาท์พุทเสมือนที่หนึ่ง จะมีขนาดเดียวกันทั้งหมดเพื่อจำกัดขอบเขตจำนวนที่จะทำให้เกิดการเรียงลำดับที่ผิดอย่างไม่มีขอบเขตที่จะเข้าไปสะสมในคิวเอาท์พุทเสมือนที่สอง

3. คิวเอาท์พุทเสมือนที่หนึ่งจะให้บริการในลักษณะที่เป็นลำดับที่กำหนดโดยสวิทช์ขั้นตอนที่หนึ่ง และเมื่อถึงรอบ แพ็กเก็ตจะออกจากคิวเอาท์พุทเสมือนที่หนึ่งและผ่านสวิทช์ขั้นตอนที่หนึ่งออกไป

4. หลังจากที่ออกจากสวิทช์ขั้นตอนที่หนึ่งแล้ว แพ็กเก็ตจะถูกเก็บไว้ในคิวเอาท์พุทเสมือนที่สองอินพุทของคิวเอาท์พุทเสมือนที่สอง คืออินพุทภายในและเอาท์พุทคือเอาท์พุทภายนอก

5. คิวเอาท์พุทเสมือนที่สองได้รับบริการในลักษณะที่เป็นลำดับที่กำหนดโดยสวิทช์ขั้นตอนที่สองและเมื่อถึงรอบ แพ็กเก็ตจะออกจากคิวเอาท์พุทเสมือนที่สองผ่านสวิทช์ขั้นตอนที่สองออกไป

6. หลังออกจากสวิทช์ขั้นตอนที่สอง แพ็กเก็ตจะรวมตัวกลับคืนในบัฟเฟอร์สำหรับรวมตัวกัน ดังนั้นคิวทั้งหมดที่มีแพ็กเก็ตจากอินพุทภายนอกที่กำหนดจะอยู่ติดกัน

7. แพ็กเก็ตจะออกจากบัฟเฟอร์สำหรับรวมตัวและออกไปโดยผ่านเอาท์พุทภายนอก

คุณสมบัติของสวิทช์ดังต่อไปนี้พบว่าเป็นประโยชน์ในงานวิจัยนี้และได้พิสูจน์ไว้ใน Chang et al.(2002)

คุณสมบัติ 1: ถ้าแพ็กเก็ตมาถึงสวิทช์ ณ เวลา t แพ็กเก็ตจะมาถึงคิวเอาท์พุทเสมือนที่สองไม่เร็วกว่า t และไม่ช้ากว่า $t + N^2$

คุณสมบัติ 2: แพ็กเก็ตที่เรียงตัวผิดลำดับได้ถูกจำกัดขอบเขตไว้ที่ $2N^2+N$

เป็นที่สังเกตว่ามีความเป็นไปได้ที่จะเพิ่มบัฟเฟอร์ที่มีขนาดจำกัดสำหรับจัดเรียงแพ็กเก็ตหลังออกจากสวิทช์ให้แก่แต่ละ เอาท์พุทภายนอก

คุณสมบัติ 3: ค่าหน่วงของแพ็กเก็ตใน EDF ได้ถูกจำกัดขอบเขตที่ผลรวมค่าหน่วงของแพ็กเก็ตในคิวเอาท์พุทของสวิทช์ FCFS และค่าคงที่ ซึ่งค่าคงที่นี้มีค่าเท่ากับ $2N^2 + N$

กลไกแบบ ASF-RR (Adaptation-Size Frames-RR with 3DQs) ร่วมกับคิวสามมิติ

ความเป็นมา

ASF-RR ร่วมกับคิวสามมิติเป็นขั้นตอนวิธีที่พัฒนาขึ้นสำหรับป้องกันการจัดเรียงตัวใหม่ของแพ็กเก็ตโดยใช้หลักการป้องกันการจัดเรียงตัวใหม่ตลอดเส้นทางภายในเราเตอร์

สมมติให้ขั้นตอนวิธีนี้ทำงานกับเฟรม ซึ่งหนึ่งเฟรมประกอบไปด้วย N แพ็กเก็ตเรียกว่าเฟรมเต็ม ถ้าเฟรมใดที่มีแพ็กเก็ตไม่ครบจำนวน N จะเติมแพ็กเก็ตว่างที่มีขนาดเดียวกันเข้าไปยังเฟรมที่ไม่เต็มนั้นจนกระทั่งเฟรมนั้นเต็ม และจะให้บริการแบบวนรอบตามลำดับเป็นที่สังเกตว่าค่าหน่วงเฉลี่ยเป็นอิสระจากลำดับที่เฟรมจะได้รับการบริการและสำหรับการจราจรที่หนาแน่นอย่างมาก (ตัวอย่างเช่น แพ็กเก็ตที่เข้ามานั้นเต็มเฟรมอยู่แล้วไม่จำเป็นต้องเติมแพ็กเก็ตว่าง) ASF-RR มีค่าหน่วงเฉลี่ยเช่นเดียวกับคิวเอาท์พุทแบบ FCFS (OQ-FCFS) ถ้ามีในคิว อย่างน้อยหนึ่งเฟรมที่เต็มแล้ว ASF-RR จะไม่ว่าง ดังนั้น ASF-RR จะอนุรักษ์การทำงานสำหรับเฟรมที่เต็ม (work conserving for full frames) ในความรู้สึกที่ว่าถ้ามีเหลืออยู่ในคิวอย่างน้อยหนึ่งเฟรมเต็มแล้วจะมีอย่างน้อยหนึ่งเฟรมเต็มที่ได้รับบริการ

การรวมกันของ ASF-RR และคิวสามมิติ

ในงานวิจัยนี้ ขั้นตอนวิธี ASF-RR ถูกนำมาใช้ร่วมกันกับคิวสามมิติในสวิทช์ขั้นตอนที่สอง พิจารณาเอาท์พุทภายนอก k กำหนดรอบ (cycle) ให้เป็นเซตของสล็อทเวลาที่ต่อเนื่องกันตามลำดับในขณะที่เอาท์พุทภายนอก k ได้รับแพ็กเก็ตเรียบร้อย

แล้วจากอินพุทภายในตั้งแต่อินพุทที่ *1* ถึง N และกำหนดเซตของคิวสามมิติ สำหรับ (i, k) ดังนี้ $\{(i,1,k), (i,2,k),\ldots,(i,N,k)\}$

สมมติให้แพ็กเก็ตสุดท้ายที่ได้รับการบริการในเซตของคิวสามมิติ มาจากอินพุทภายใน j_{last} เพราะว่าคุณสมบัติของตัวถ่วงสมดุลย์ คือลำดับถัดไปของแพ็กเก็ตของสายข้อมูล F_{ik} จะมาจาก อินพุทภายในที่ $j_{las}+ 1 \ (mod\ N)$ ให้P_{ik} เป็นพอยน์เตอร์ที่ชี้ไปยังแพ็กเก็ตถัดไปที่ถูกต้องตามลำดับ $p_{ik} \leftarrow j_{last} + 1 (mod\ N)$ กำหนดเฟรมสำหรับ(i,k) ดังนี้ $f(i,k) = \{(i, P_{ik}, k)$ $,(i, P_{ik}+ 1,k), \ (i,N,k)\}$ และเฟรม $f(i,k)$ เต็ม ถ้าทุก คิวสามมิติ (i,j,k) สำหรับ $j = \{P_{ik},\ldots,N\}$ ไม่ว่าง ถ้าเฟรมเต็มดังนั้นแพ็กเก็ตตามลำดับถัดไปจะอยู่ในคิวสามมิติ (i, P_{ik}, k) และถัดไปจะอยู่ใน $(i, P_{ik}+ 1,k)$ เป็นเช่นนี้เรื่อยไปจนกระทั่งถึงแพ็กเก็ตลำดับ(i,N,k) สมมติว่าขั้นตอนวิธี EDF ใช้โครงสร้างคิวสามมิติ ดังนั้น ณ การเริ่มต้นของทุกรอบจะมีการจัดเรียงตัวของคิวสามมิติใหม่ คิวทั้งหมดจะมีแพ็กเก็ตที่มาจากอินพุทภายนอกที่กำหนดอยู่ติดกัน หลักการนี้คือกุญแจสำคัญในการป้องกันและจำกัดขอบเขตแพ็กเก็ตที่อยู่ในเฟรมไม่ให้ผิดลำดับ

ใน ASF-RR ร่วมกับคิวสามมิตินั้น เอาท์พุทภายนอกจะอ่านแพ็กเก็ตทั้งหมดในเฟรมที่เต็มจากอินพุทภายนอก ก่อนที่จะย้ายไปอ่านเฟรมที่เต็มจากอินพุทภายนอกถัดไป เอาท์พุทภายนอก k ใช้พอยน์เตอร์ $P_{ff}(k)$ แบบวนรอบในการจำว่าเฟรมเต็มนี้มาจากอินพุทภายนอกตัวใด

เพื่อความถูกต้องในการคำนวณจะพิจารณาที่เอาท์พุทภายนอก k ณ จุดเริ่มต้นของทุกๆรอบ โดยการเริ่มต้นที่ $P_{ff}(k)$จะหาเฟรมถัดไปที่มาถึงจาก อินพุทภายนอกแล้วปรับปรุง $p_{ff}(k) \leftarrow i_{ff} + 1 (mod\ N)$

ประสิทธิภาพของ ASF-RR ร่วมกับคิวสามมิติ

ในส่วนนี้จะแสดงว่าค่าหน่วงเฉลี่ยสำหรับขั้นตอนวิธี ASF-RR ร่วมกับคิวสามมิติ มีค่าต่ำกว่าค่าหน่วงเฉลี่ยของคิวเอาท์พุทรวมกับค่าคงที่และ ASF-RR ร่วมกับคิวสามมิติ มีค่าประสิทธิผลเช่นเดียวกับคิวเอาท์พุท การพิสูจน์ใช้หลักการที่ว่า ASF-RR ร่วมกับคิวสามมิติ ทำงานแบบอนุรักษ์การทำงานสำหรับเฟรมที่เต็ม

คำจำกัดความ

พิจารณาเฉพาะแพ็กเก็ตที่มีปลายทางไปยังเอาท์พุทภายนอก k ที่กำหนด ค่าต่อไปนี้ถูกกำหนดเพื่อใช้ในการวิเคราะห์ ดังแสดงในภาพที่ 4

1. $A_i^*(t)$ เป็นจำนวนสะสมของแพ็กเก็ตที่จะไปยัง เอาท์พุทภายนอก k ซึ่งบัดนี้มาถึง อินพุทภายนอก i แล้ว จนถึงเวลา t รวมสล็อทเวลา t ด้วย และเป็นดัชนีของแพ็กเก็ตสุดท้ายจาก F_{ik}ที่มาถึงอินพุทภายนอก i

2. $A^o(t) = \sum_{i=1}^{N} A_i^*(t)$ เป็นจำนวนทั้งหมดของแพ็กเก็ตที่มีปลายทางไปยัง k ซึ่งมาถึงสวิทช์จนกระทั่งถึงเวลา t

3. $A_i(t)$ เป็นดัชนีของแพ็กเก็ตสุดท้ายในสายข้อมูล F_{ik} ที่มาถึง อินพุทภายใน โดยการเริ่มต้นของรอบปัจจุบันโดยไม่มีแพ็กเก็ตใดผิดไปจากลำดับของสายข้อมูลที่เข้ามา ไม่มีแพ็กเก็ตใดที่อยู่ข้างหน้าแพ็กเก็ตนี้ที่ผิดลำดับ เรียกว่า แพ็กเก็ตลำดับ (ordered cells) จำนวนทั้งหมดของแพ็กเก็ตลำดับคือ $A(t) = \sum_{i=1}^{N} A_i(t)$ โดยนิยาม ถ้า t' เป็นการเริ่มต้นของรอบของ เอาท์พุทภายนอก k (ซึ่ง $t' = k + 1 + l \cdot N$, l เป็น เลขจำนวนเต็ม) แล้ว $A_i(t) = A_i(t')$ เมื่อ $t' = k + 1 + l \cdot N$ ดังได้แสดงในคุณสมบัติ 1 จะได้ว่า $A_i^*(t' - N^2) \le A_i(t') \le A_i^*(t')$ (สมมติให้ค่าเหล่านี้เป็น 0 เมื่อ $t' \le 0$)

4. $B_i(t)$ เป็นจำนวนแพ็กเก็ตลำดับที่ได้รับการ
บริการ แล้วโดยสวิทช์ขั้นตอนที่สอง จนกระทั่งถึงสล็อทเวลา t และ $B(t) = \sum_{i=1}^{N} B_i(t)$

5. $q(t) = A(t) - B(t)$ เป็นจำนวนของแพ็กเก็ตลำดับที่อยู่ในคิวใน คิวสามมิติ ที่มีปลายทางไปที่ k

6. $A_{ff}(t) = \sum_{i=1}^{N} N.\left\lfloor \frac{A_i(t)}{N} \right\rfloor$ เป็นจำนวนของแพ็กเก็ต ในเฟรมเต็ม ที่ได้มาถึง คิวสามมิติ เรียบร้อยแล้วและ มีปลายทางไปยัง k

7. $B_{ff}(t) = \sum_{i=1}^{N} N.\left\lfloor \frac{B_i(t)}{N} \right\rfloor$ เป็นจำนวนของแพ็กเก็ตในเฟรมเต็มที่ได้รับการบริการเรียบร้อยแล้วโดยสวิทช์ขั้นตอนที่สอง

8. จำนวนของเฟรมเต็มใน คิวสามมิติ คือ

$$ff(t) = \frac{A_{ff}(t)}{N} - \frac{B_{ff}(t)}{N} = \sum_{i=1}^{N} \left\lfloor \frac{A_i(t)}{N} \right\rfloor - \left\lfloor \frac{B_i(t)}{N} \right\rfloor$$

และ $q_{ff}(t) = N \cdot ff(t) = A_{ff}(t) - B_{ff}(t)$

9. $B^{oq}(t)$ และ $q^{oq}(t)$ เป็นจำนวนสะสมของแพ็กเก็ตที่ได้รับบริการและความยาวของคิวตามลำดับในสวิทช์ คิวเอาท์พุท (queue output switch) สปีดอัพหนึ่งซึ่งจำนวนสะสมของการมาถึงคือ $A^o(t)$

10. $B_D^{OQ}(t)$ และ $q_D^{OQ}(t)$ เป็นจำนวนสะสมของแพ็กเก็ตที่ได้รับการบริการและความยาวของคิวตามลำดับในสวิทช์ คิวเอาท์พุทแบบหน่วง (delayed output queue switch) หรืออาจกล่าวได้ว่าสวิทช์คิวเอาท์พุทแบบหน่วงสามารถแทนสวิทช์คิวเอาท์พุทสปีดอัพหนึ่งที่มีลำดับการมาถึงเช่นเดียวกับที่อินพุทภายในของสวิทช์ขั้นตอนที่สอง คุณสมบัติที่รู้จักแพร่หลายของคิวเอาท์พุทมีดังต่อไปนี้

$$q_D^{OQ}(t) = A(t) - B_p^{\infty}(t),$$

$$q_D^{OQ}(t) = \max_{0\le s\le t} \left(A(t) - A(s) - (t-s)\right),$$

$$B_D^{OQ}(t) = \max_{0\le s\le t} \left(A(s) + (t-s)\right), \text{ และ}$$

$$B_D^{OQ}(t) - B_D^{OQ}(s) \le t - s, \text{เมื่อ} \qquad s \le t$$

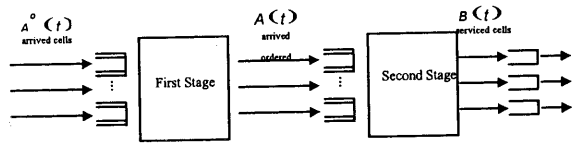
Keslassy and McKeown (2002)

11. $t' = k + l + l \cdot N$ เป็นสล็อทเวลาใดๆ เมื่อรอบของเอาท์พุทภายนอก เริ่มต้น l เป็นเลขจำนวนเต็มใดๆ

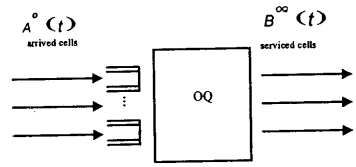
ค่าหน่วงเฉลี่ยของ ASF-RR ร่วมกับคิวสามมิติอยู่ภายในค่าคงที่ของคิวเอาท์พุท

ในส่วนนี้จะแสดงให้เห็นว่าค่าหน่วงเฉลี่ยสำหรับ ASF-RR ร่วมกับคิวสามมิติ อยู่ภายในค่าหน่วงที่เป็นค่าคงที่ของค่าหน่วงเฉลี่ยของสวิทช์คิวเอาท์พุท สำหรับการจราจรที่เข้ามาเป็นแบบเดียวกัน

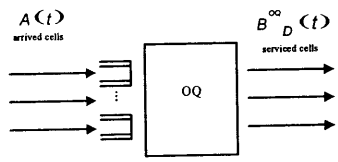
เปรียบเทียบ ASF-RR ร่วมกับคิวสามมิติและคิวเอาท์พุทแบบหน่วงเป็นอันดับแรก ซึ่งคิวเอาท์พุทแบบหน่วงเป็นคิวเอาท์พุทที่มีลำดับการเข้ามาของแพ็กเก็ตเป็นแบบเดียวกับสวิทช์ขั้นตอนที่สอง จะแสดงว่า ASF-RR ร่วมกับคิวสามมิติ เป็นการอนุรักษ์การทำงานสำหรับเฟรมที่เต็ม ทำให้เกิดการจำกัดขอบเขตค่าหน่วงเฉลี่ยที่ต่างจากแบบจำลอง



(ก) สวิทช์สองขั้นตอนกับ AFS-RR ร่วมกับคิวแบบ สามมิติ



(ข) แบบจำลอง OQ สำหรับสวิทช์ขั้นตอนแรก (กับการมาถึงของ $A^{\circ}(t)$



(ค) แบบจำลอง OQ สำหรับสวิทช์ขั้นตอนสอง (กับการมาถึงของ $A(t)$)

ภาพที่ 4 ความหมายของเทอมต่างๆ

คิวเอาท์พุทแบบหน่วง แล้วเปรียบเทียบแบบจำลองคิวเอาท์พุทแบบหน่วงกับสวิทช์คิวเอาท์พุท หรือเรียกว่าสวิทช์คิวเอาท์พุทแบบธรรมดา ซึ่งสวิทช์คิวเอาท์พุทนี้มีลำดับการมาถึงของแพ็กเก็ตเช่นเดียวกับ สวิทช์ขั้นตอนที่หนึ่ง โดยการใช้ค่าหน่วงในการจำกัดขอบเขตค่าหน่วงเฉลี่ยที่ต่างกันระหว่าง ASF-RR ร่วมกับคิวสามมิติกับสวิทช์คิวเอาท์พุทเริ่มต้นเมื่อมีอย่างน้อยหนึ่งเฟรมเข้ามาที่คิว

สามมิติ จำนวนเฟรมที่ได้รับการบริการจะเพิ่มขึ้นทีละ หนึ่งเฟรมในรอบถัดไป

บทตั้งและทฤษฎีต่อไปนี้ได้พิสูจน์ไว้ใน Chiewthanakul and Waiyanon (2006)

บทตั้ง 1: ถ้ามีอย่างน้อยหนึ่งเฟรมมาถึงสวิทช์ขั้นตอนที่สอง ดังนั้น $B_{ff}(t'+N) = B_{ff}(t') + N$

บทตั้ง 2: $B_{ff}(t') \geq B_D^{QQ}(t') - N^2$

บทตั้ง 3: $q_{ff}(t') - q_D^{QQ}(t') \leq N^2 + N - 1$ และ $q(t') - q_D^{QQ}(t') \leq N^2$

บทตั้ง 4: ถ้ามีอย่างน้อยหนึ่งเฟรมมาถึงสวิทช์ขั้นตอนที่สอง ดังนั้น $B_{ff}\left(t' + q_{ff}(t')\right) = A_{ff}(t')$

บทตั้ง 5: $B\left(t' + q_{ff}(t') + N(2N-2)\right) \geq A(t')$

ทฤษฎี 1: ค่าหน่วงเฉลี่ยสำหรับ ASF-RR ร่วมกับคิวสามมิติมีค่าน้อยกว่าค่าหน่วงเฉลี่ยสำหรับคิวเอาท์พุทแบบหน่วงรวมกับค่าคงที่ $3N^2 - 2$

ทฤษฎี 2: ค่าหน่วงเฉลี่ยของ ASF-RR ร่วมกับคิวสามมิติ มีค่าน้อยกว่าค่าหน่วงเฉลี่ยของคิวเอาท์พุทรวมกับค่าคงที่ $4N^2 + N - 3$

ASF-RR ร่วมกับคิวสามมิติมีค่าประสิทธิผลเช่นเดียวกับคิวเอาท์พุท

พิจารณาสวิทช์ที่มีอัตราเข้ามาของแพ็กเก็ตเป็น ρ_{ik} และขนาดของคิวทั้งหมดเป็น $Q(t)$ เมื่อ t เป็นสล็อทเวลาในปัจจุบัน

1. โหลดของการจราจรที่เข้ามาเป็น

$$\rho = max\left(\max_k\left(\sum_{i=1}^{N}\rho_{ik}\right), \max_i\left(\sum_{k=1}^{N}\rho_{ik}\right)\right)$$

การจราจรที่เข้ามานั้นจะพอรับได้ถ้า $\rho < 1$

2. สวิทช์มีความคงตัวอย่างมากถ้า $\limsup_{t\to\infty}(E[Q(t)]) < \infty$ Leonardi (2001)

3. สวิทช์มีค่าประสิทธิผล 100% ถ้าสวิทช์มีความคงตัว หรือเมื่อไรก็ตามที่การจราจรที่เข้ามานั้นรับได้ และพบว่าขอบเขตของค่าหน่วงเฉลี่ยที่ต่างกันระหว่าง ASF-RR ร่วมกับคิวสามมิติและสวิทช์คิวเอาท์พุท มีอยู่จริง ลำดับต่อมา ASF-RR ร่วมกับคิวสามมิติ มีค่าประสิทธิผลเหมือน คิวเอาท์พุท ดังเช่นแสดงไว้ใน ทฤษฎี 3

ทฤษฎี 3: ASF-RR ร่วมกับคิวสามมิติและคิวเอาท์พุทมีค่าประสิทธิผลเหมือนกัน

คุณสมบัติ 4: สมมติว่ารูปแบบการจราจรที่เข้ามาจากอินพุทภายนอก i ไปยัง เอาท์พุทภายนอก k เป็น (σ, ρ_{ik}) ถ้าการจราจรพอที่จะรับได้ ดังนั้น ASF-RR ร่วมกับคิวสามมิติ ให้ค่าประสิทธิผล 100% พิสูจน์ใน Chiewthanakul and Waiyanon (2006)

วิธีการทดลอง

ทดลองโดยจำลองการทำงานของ ASF-RR ร่วมกับคิวสามมิติ ลงในโปรแกรม MATLAB ใช้การกระจายแบบ Geom/Geom/1 ให้การเข้ามาของข้อมูลในแต่ละสล็อทเวลาด้วยความน่าจะเป็นเท่ากับ p จากนั้นจำลองเพื่อหาค่าหน่วงเฉลี่ยของระบบที่ค่าโหลดแฟคเตอร์ (load factor: ค่าอัตราส่วนของความน่าจะเป็นในการเกิดของแพ็กเก็ตในคิวกับความน่าจะเป็นในการดับของแพ็กเก็ตในคิว) ต่าง ๆ ที่สภาวะคงตัว (steady state)

ในการศึกษานี้จะใช้วิธีการประเมินผลการทำงานของขั้นตอนวิธี ASF-RR ร่วมกับคิวสามมิติ และ FFF ด้วยการหาประสิทธิภาพโดยการหาค่าหน่วงเฉลี่ยที่ ค่าโหลดแฟคเตอร์ต่าง ๆ

ในงานวิจัยนี้ได้แบ่งการทดลองออกเป็น 3 ชุด ด้วยกันคือ

ชุดที่ 1. ทดสอบเปรียบเทียบค่าหน่วงเฉลี่ยระหว่าง ASF-RR ร่วมกับคิวสามมิติ และ FFF

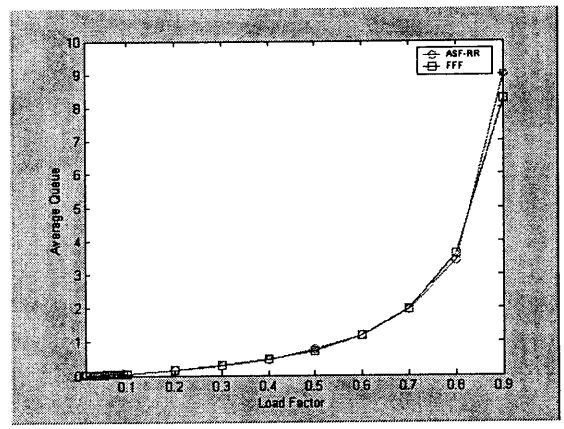
ชุดที่ 2. ทดสอบเปรียบเทียบจำนวนแพ็กเก็ตที่ค้างในคิวระหว่าง ASF-RR ร่วมกับคิวสามมิติ และ FFF

ชุดที่ 3. หาค่าหน่วงเฉลี่ย ณ ค่าโหลดแฟคเตอร์ต่างๆ กันที่ขนาดของแพ็กเก็ตต่างๆ

ผลการทดลอง

ในส่วนนี้จะนำเสนอผลการจำลองซึ่งแสดงถึงค่าหน่วงเฉลี่ยของวิธีการที่ได้นำเสนอและเปรียบเทียบกับวิธี FFF โดยกำหนดให้การเข้ามาของข้อมูลมีการแจกแจงแบบเรขาคณิต ด้วยพารามิเตอร์ p เท่ากับ 0.5 ข้อมูลที่มีความยาวแบบแปรผันถูกแบ่งออกให้มีขนาดคงที่

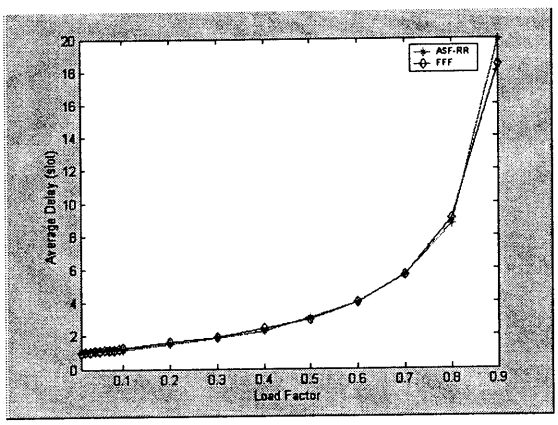
ภาพที่ 5 แสดงความสัมพันธ์ระหว่างค่าเฉลี่ยของคิวกับค่าโหลดแฟคเตอร์ ของ FFF และ ASF-RR ร่วมกับคิวสามมิติ พิจารณาที่ช่วงโหลดแฟคเตอร์ มีค่าน้อยกว่า 0.5 จะเห็นว่าค่าเฉลี่ยของคิวจะต่ำ คือมีค่าน้อยกว่า 1 แพ็กเก็ต โดยทั้ง FFF และ ASF-RR ร่วมกับคิวสามมิติมีค่าเฉลี่ยของคิวที่ใกล้เคียงกันมากแต่เมื่อโหลดแฟคเตอร์ มีค่ามากกว่า 0.5 ทั้ง FFF และ ASF-RR ร่วมกับคิวสามมิติ มีค่าเฉลี่ยของคิวที่แตกต่างกันเพียงเล็กน้อย เมื่อ โหลดแฟคเตอร์ มีค่ามากขึ้น ค่าเฉลี่ยของคิวทั้งใน FFF และASF-RR ร่วมกับคิวสามมิติ จะมีการเพิ่มขึ้นอย่างรวดเร็ว จากกราฟยังพบว่าค่าเฉลี่ยของคิวถึงแม้ว่า โหลดแฟคเตอร์มีค่าเข้าใกล้ 1 ก็ยังคงมีค่าน้อยกว่า 10 ซึ่งถือว่ามีค่าต่ำมาก ดังนั้นค่าเฉลี่ยของคิวจากทั้ง FFFและ ASF-RR



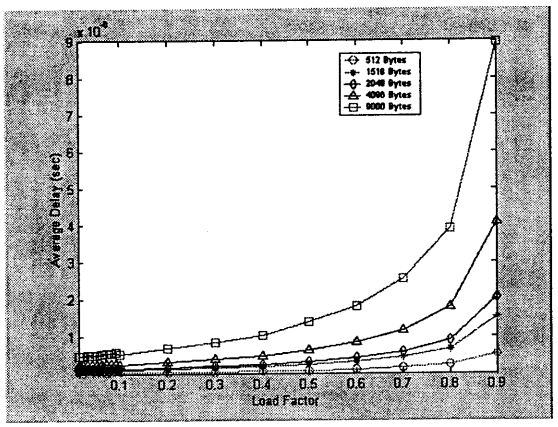
ภาพที่ 5 ความสัมพันธ์ระหว่างค่าเฉลี่ยของ คิวกับค่าโหลดแฟคเตอร์

ร่วมกับคิวสามมิติมีค่าที่ใกล้เคียงกันมาก

จากภาพที่ 6 แสดงความสัมพันธ์ระหว่างค่าหน่วงเฉลี่ย กับค่าโหลดแฟคเตอร์ ของ FFF และ ASF-RR ร่วมกับคิวสามมิติ จะเห็นว่ามีลักษณะเช่นเดียวกับภาพที่ 5 พิจารณาที่ช่วง โหลดแฟคเตอร์ มีค่าน้อยกว่า 0.5 จะเห็นว่าค่าหน่วงเฉลี่ยจะต่ำ คือ มีค่าน้อยกว่า 4 สล็อท โดยทั้ง FFF และ ASF-RRร่วมกับคิวสามมิติ มีค่าหน่วงเฉลี่ยที่ใกล้เคียงกันมากแต่เมื่อโหลดแฟคเตอร์ มีค่ามากกว่า 0.5 ทั้ง FFFและ ASF-RR ร่วมกับคิวสามมิติ มีค่าหน่วงเฉลี่ยที่แตกต่างกันเพียงเล็กน้อย เมื่อโหลดแฟคเตอร์ มีค่ามากขึ้นจะเห็นว่าค่าหน่วงเฉลี่ยเพิ่มสูงขึ้นอย่างรวดเร็ว จากกราฟยังพบว่าถึงแม้ว่าโหลดแฟคเตอร์มีค่าเป็น 0.9 ค่าหน่วงเฉลี่ยมีค่าประมาณ 20 สล็อท ซึ่งถือว่ามีค่าต่ำ ดังนั้นค่าหน่วงเฉลี่ยจากทั้ง FFF และ ASF-RR ร่วมกับคิวสามมิติมีค่าที่ใกล้เคียงกันมาก



ภาพที่ 6 ความสัมพันธ์ระหว่างค่าความหน่วงเฉลี่ยกับค่าโหลดแฟคเตอร์



ภาพที่ 7 ความสัมพันธ์ระหว่างค่าหน่วงเฉลี่ยกับค่าโหลดแฟคเตอร์ที่ขนาดแพ็กเก็ตต่างกัน

จากภาพที่ 7 แสดงความสัมพันธ์ระหว่างค่าหน่วงเฉลี่ยและโหลดแฟคเตอร์ โดยทดลองที่ขนาดแพ็กเก็ต 512, 1518, 2048, 4096 และ 9000 ไบต์ จะเห็นว่าเมื่อค่า โหลดแฟคเตอร์ มีค่าต่ำกว่า 0.5 ค่าหน่วงเฉลี่ยของขนาดแพ็กเก็ต 512, 1518, 2048 และ 4096 ไบต์ไม่สูงมากนัก คือมีค่าต่ำกว่า 1 ไมโครวินาที ขณะที่แพ็กเก็ตที่มีขนาด 9000 ไบต์ มีค่าหน่วงเฉลี่ยไม่เกิน 2 ไมโครวินาที แต่เมื่อค่าโหลดแฟคเตอร์มีค่ามากกว่า 0.5 ค่าหน่วงเฉลี่ยของแพ็กเก็ตทุกขนาดจะมีค่าสูงขึ้นโดยเฉพาะเมื่อโหลดแฟคเตอร์มีค่าเข้าใกล้ 1 ค่าหน่วงเฉลี่ยจะเพิ่มสูงขึ้นอย่างเห็นได้ชัด แพ็กเก็ตขนาด 512 ไบต์ จะมีค่าหน่วงเฉลี่ยต่ำสุด ในขณะที่แพ็กเก็ตขนาด 1518, 2048, 4096 ไบต์จะมีค่าหน่วงเฉลี่ยเพิ่มมากขึ้นตามลำดับ และแพ็กเก็ตขนาด 9000 ไบต์จะมีค่าหน่วงเฉลี่ยสูงที่สุด แสดงว่าเมื่อแพ็กเก็ตมีขนาดมากขึ้นจะมีค่าหน่วงเฉลี่ยมากขึ้นด้วย ดังนั้นจากกราฟจะเห็นว่าเมื่อโหลดแฟคเตอร์มีขนาดต่ำๆ ขนาดของแพ็กเก็ตที่ควรเลือกใช้คือ 2048 ไบต์ ในขณะที่โหลดแฟคเตอร์มีค่าสูง ขนาดแพ็กเก็ตที่ควรเลือกใช้คือ 512 ไบต์

บทสรุป

จากการศึกษาประสิทธิภาพในการป้องกันการจัดเรียงตัวใหม่ของแพ็กเก็ตในเราเตอร์แบบถ่วงสมดุลย์ โดยจำลองการทำงานบนระบบเครือข่าย ซึ่งได้นำเสนอขั้นตอนวิธี ASF-RR ร่วมกับคิวสามมิติ และเปรียบเทียบประสิทธิภาพกับ FFF โดยนำเสนอผลการวิเคราะห์ในรูปของค่าหน่วงเฉลี่ย ซึ่งจากผลการจำลองแสดงให้เห็นว่าขั้นตอนวิธีที่ได้นำเสนอนั้นให้ค่าหน่วงเฉลี่ยที่ใกล้เคียงกับ FFF สำหรับประสิทธิภาพในเทอมของประสิทธิผลนั้น ASF-RR ร่วมกับคิวสามมิติ สามารถรับประกันค่าประสิทธิผล 100 % Chiewthanakul and Waiyanon (2006) นอกจากนี้ ขั้นตอนวิธี ASF-RR ร่วมกับคิวสามมิติมีข้อดีคือสามารถแก้ปัญหาในเรื่องของการขัดขวางแพ็กเก็ตหัวแถว ให้ค่าประสิทธิผลที่สูง รับประกันค่าหน่วง และสามารถแก้ปัญหาการสะสมของแพ็กเก็ตที่ผิดลำดับในบัฟเฟอร์ที่เกิดขึ้นใน FFF ได้

เอกสารอ้างอิง

Anderson, T. E., Owicki, S. S., Saxe, J. B. and Walrand, J. 1993. High speed switch scheduling for local area networks. ACM Trans. On Computer Systems. Vol. 11, No.4 :319-352.

Blanton, E. and Allman, M. 2002. On making TCP more robust to packet reordering. ACM Computer Communication Review. Vol. 32, No. 1: 20-30.

Chang, C. S., Lee, D. S. and Shih, Y. J. 2004. Mailbox switch: a scalable two-stage switch architecture for conflict resolution of ordered packets. IEEE Infocom'04. Hong Kong.

Chang, C. S., Lee, D. S. and Yue, C. Y. 2003. Providing guaranteed rate services in the load balanced Birkhoff-von Neumann switches. IEEE Infocom'03.

Chang, C. S., Lee, D. S. and Lien, C. M. 2002. Load balanced Birkhoff-von Neumann switches, Part II: multi-stage buffering. Computer Communications. Vol. 25, No. 6: 623-634.

Chiewthanakul, B.and Waiyanon, K. 2006. The Performance on the Prevention of Packet Reordering in load-balanced Router. Proceedings of the Technology and Innovation for Sustainable Development Conference (TISD). Khon Kaen, Thailand.843-847

Dai, J. G. and Praphakar, B. 2000.The throughput of data switches with and without speedup. IEEE Infocom'00. Tel Aviv, Israel. Vol. 2: 556-564.

Fomenkov, M., Keys, K., Moore, D. and Claffy, K. 2004. A longitudinal study of internet traffic from 1998-2001: a view from 20 high performance sites. Proc. Of WISICT'04. Cancun, Mexico.

Keslassy, I., Chang, C. S., McKeown, N. and Lee, D.S. 2005. Optimal load-balancing. IEEE Infocom'05.

Keslassy, I. Chuang., S. T. and McKeown, N. 2004. A load-balanced switch with an arbitrary number of linecards. IEEE Infocom'04. Hong Kong.

Keslassy, I., Chuang, S. T. Yu, K. Miller, D. Horowitz, M. Solgaard ,O. and McKeown, N. 2003. Scaling internet routers using optics. *ACM SIGCOM'03*. Karlsruhe, Germany.

Keslassy, I.,and McKeown, N. 2002. Maintaining packet order in two-stage switches. IEEE Infocom'02.

Leonardi, E., Mellia, M., Neri, F. and Marsan, M. A. 2001. On the stability of input-queued switches with speed-up. IEEE/ACM Trans. on Networking. Vol. 9, No. 1:104-118.

Marsan, M. A., Bianco,A., Giaccone, P E., Leonardi,E. and Neri ,F. 2001. Packet scheduling in input-queued cell-based switches. IEEE Infocom'01. Anchorage, Alaska.

McKeown, N. 1999. iSLIP: A scheduling algorithm for input-queued switches. IEEE Trans. on Networking, Vol. 7, No.2.

McKeown, N., Mekkittikul, A., Anantharam V. and Walrand, J. 1999. Achieving 100% throughput in an input-queued switch. IEEE Trans. on Communications. Vol. 47, No. 8.

Valiant, L. G. 1982. A scheme for fast parallel communication. SIAM Journal on Computing. Vol. 11, No.2:350-361.