マルチメディア通信と分散処理　62－6
（1993. 9. 20）

Ｂ－ＩＳＤＮのためのインテリジェントセルフルーティングアルゴリズム

ラシッド　イマド、小林広明、中村維男

東北大学工学部機械知能工学科
〒９８０　仙台市荒巻字青葉

あらまし　本論文は、非同期転送方式（ＡＴＭ）によるＢ－ＩＳＤＮのための新しいインテリジェントセルフルーティングアルゴリズムを提案する。本アルゴリズムは、ants routing とよぶ新しい輻輳制御方式に基づいている。本アルゴリズムにより、ネットワークのスイッチの出力ポート上にバースト的に発生するトラフィックを最適経路制御し、その結果、高いスループットと低パケット損失率を達成することができる。ants routing が必要とする各スイッチの輻輳状況は、常にモニタされ、そして隣接スイッチ間で共有される。さらに本論文では、待ち行列モデルに基づいた解析モデルにより、本方式の有効性について議論する。

和文キーワード　Ｂ－ＩＳＤＮ、インテリジェントスイッチ、輻輳制御、セルフルーティング

An Intelligent Self-Routing Algorithm for B-ISDN

Emad Rashid　Hiroaki Kobayashi,　and Tadao Nakamura

Department of Machine Intelligence and
System Engineering,　Faculty of Engineering
Tohoku University, Aramaki Aza Aoba, Aoba-Ku　Sendai 980, JAPAN

Abstract　This paper presents a new self-routing algorithm for broadband ISDN's asynchronous transfer mode (ATM) switching networks. The routing algorithm is ambuscade in a switch for congestion control called Ants Routing. The congestion is controlled through regulating the input traffic rate to the switch element that has congestion on one of its output p ports. high throughput and low packet loss probability can be achieved by rerouting packets' arrival due to the presence of bursty traffic on a switch's output port. The rerouting algorithm is based on the information of congestion status of each switch, which can be distributed among neighboring switches. Mathematical analysis based on the queuing model shows that our algorithm has capability of congestion avoidance on the interconnection network and packet loss improvement, especially when traffic is bursty.

英文 key words　B-ISDN, Ants Routing, Intelligent Switch, Congestion control

1. INTRODUCTION

ATM is expected to be a target transfer mode for the broadband Integrated Service Digital Network (B-ISDN) [1] [2]. An ATM network is a multimedia network that handles various traffic rates such as video and data applications. These applications require high-speed switching networks to transfer large data rates from source to destination [3]. Therefore, these networks may prone to congestion, because of the multimedia communication in which the aggregate number of packets is characterized in a variable rate transmission and the number of packets arriving at each link exceeds switch's buffer size (resources) [2]. Besides, the packets will continue arriving at the switch even there is congestion in the network. This phenomenon leads to a high packet loss probability inside the network and to a low network efficiency. To achieve a high throughput and a low packet loss probability, it is required to reconsider the design of the interconnection network architecture, the routing algorithms, and the design of the switch element itself to be an intelligent switch element. These issues will be discussed in the following sections.

Many research efforts have been devoting to solve the congestion problem in multimedia networks [4] [5], and rerouting packets to attain high throughput, especially for nonuniform traffic [3] [5]. Many research efforts in congestion control have been focusing on regulating the input traffic using techniques such as distributed source control, channel-sharing [3],[13],and Knockout Switch [10]. In this paper, we propose a new rerouting algorithm for self-routing switches, under network congestion condition and independent of the network routing algorithm during the call setup. The congestion is controlled through regulating the input traffic rate to the switch element that has congestion on one of its output ports. The control mechanism is achieved by antecedent rerouting some of the packets that have addresses destined for a high traffic link. This new routing technique, which we called it *Ants Routing*, is based on the information
that has been collected in the switch element concerning about the traffic status of the input/output ports. Each node (switch element) shares its information about the traffic levels with other adjacent nodes. The term Ant routing comes from the analogy of ants changing its way or route when they receive messages from other ants to discover a new object, but the previous route remains working. In this paper, self-rerouting for intelligent switch is proposed. Then, we evaluate the performance of the *Ants routing* scheme in view of the probability of packet loss. The results demonstrate that the *Ants routing* scheme provides significant advantages of relieving the network of the congestion and reducing the consequence of the switch's buffer size. However, using this scheme, we show how to engineer various parameters used to satisfy requirements of packets loss inside interconnection network.

2. SWITCH ARCHITECTURE

2-1. Structure of Switch Element

The most important thinks need to be emphasized in the switch design; the first is the internal fabric used in which the switches able to route the packets from any of its input ports to any of its output ports without blocking. The blocking characteristic has a serious effect to limit the throughput of the switch element [12]. The second is a low communication latency, this property depends on the type of switching technology used to transfer packets from input ports to output ports [11]. The third is intelligent packet switch routing, and the switch element is required to be intelligent to avoid interconnection fabric congestion and to gain high throughput. This new capability for the switch is based on a software embedded in its structure to be added to the switch routing algorithms. Each switch element has to share the information with adjacent elements about congestion level in interconnection fabric. This information uses to reroute the packets which have address destined for a high traffic link, to take other lower traffic links. The new routing algorithm has no action when the traffic in the interconnection network is in a normal traffic fashion.

2-2. Interconnection switch network

One of the desirable requirements for high throughput and low packet loss probability in ATM network is a network topology which is a key factor in determining a suitable architecture, and many topologies have been considered for interconnection switch networks [6]. Each architecture has its own advantages and disadvantages. Many proposed switching systems employed nonblocking interconnection network [7], to avoid switch fabric congestion and to minimize the switch fabric complexity [8], and to provide scalable interconnection switches for less loss packet and high throughput performance [9] [10]. Fig. 1 shows the proposed large scale ATM switch architecture, that is a *NxN* a complete nonblocking interconnection structure of the four-stage Clos network[8]. Each switch element has the identical number of input/output ports.

2-3. Routing Algorithm

In this paper, we propose nonblocking crossbar switch element which can route the packets from any of its input ports to any other output ports. Therefore, this technique may occur many of packet arrivals to the input ports perhaps that will be routed to the same output port. The aggregation of the packets that chooses a specific output may case congestion on this port because of the bursty traffic pattern that exceeds buffer size of the link in case of nonuniform traffic rate, e.g., to transmitting video images that use a large data rate.

This phenomenon is occurred by many reasons. The most important reason is a weakness of routing algorithm during the call setup to select the destination. Therefore, the packet loss probability will be high, and the switch performance will be low. We can resolve such a problem and to relieve congestion in the switch network by reducing the traffic on the switch that has congestion in one of its output port. Congestion control scheme is performed by rerouting algorithm called *Ants routing* applied to a particular interconnection switching network, we will discuss this routing in detail in the next section. The congestion control mechanism leads to a lower packet loss probability, and a higher utilization of network resources by deflecting some packet arrivals to choose a route that has low traffic rate.

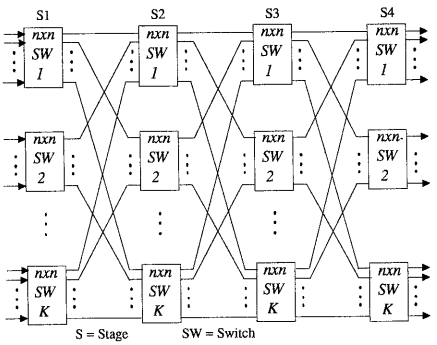

Fig. 1. The proposed switch architecture

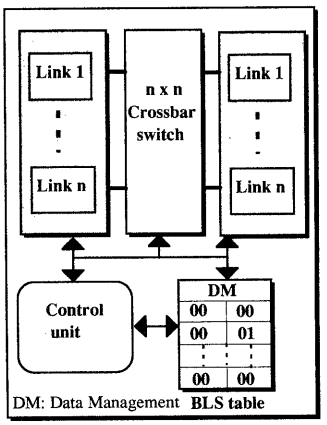

Fig. 2. Switch element architecture

3. THE SOURCE RATE SWITCH REDUCTION SCHEME

We consider each switch element has identical input/output ports size. To construct self-routing interconnection ATM switches, a multi-stage non-blocking interconnection network is proposed. The first three stages perform as self-routing interconnection fabric. All switches are connected to each others through control links to transfer information about links' traffic status. The switch element has two small memories, one is to store its input/output status as binary form table called *Binary link status BLS* table, the other is a control unit to control the links and to use the information in BLS-table to avoid congestion in the interconnection switch fabric. The information in the BLS table of a switch element can share with other BLS tables for adjacent switch elements. The input link or output link has three probability traffic levels (weights) which are represented in binary and stored in BLS table. The traffic on the link may be low/normal traffic level β, high traffic level ζ, or congestion level Ξ, where β, ζ, and Ξ are represented in binary as 00, 01, and 10, respectively. All this information is store in BLS-table, as in Fig. 2.

The switch element has three kinds of control; First, in a low or a normal traffic, the switch control is to translate packets arrival and route the packets to the destination address that correspond to the address in the header. Second is a high traffic, in which the link has a few available resources, the switch control in this case, is to start rerouting packets to choose alternative paths. Third is traffic congestion at one of switch's output ports, in which the link has no more available resources to accept arriving packets, the switch control has to change the information in BLS table that is related to the switch's input port of the same switch element. For these reasons why we choose three probability traffic levels which are represented in binary (00, 01, 10) and stored in BLS table.

3-1. Algorithm for Data Control in Switch's BLS table

This subsection describes the data control algorithm for the BLS's data which represents the level of the traffic at the input and output port's, also describes data control among BLS's tables in adjacent switch elements.

Hence, the control algorithm is achieved by rules of the knowledge embedded in each switch element. Obviously, there are some rules to transfer the weight from outlet to inlet in the same BLS table for the switch element. However, there are three traffic conditions that may occur on the outlet links: First, if the switch's BLS table has β weight value for outlet ports then we can say that the switch's outlet ports have low or normal traffic fashion. Therefore, there is no action that can be taken from the switch element for transferring weight to inlets ports, and then $\mathcal{B}_i^o[SW_z, S_h] = \beta$, $(i = 1, 2, \ldots, n)$, $(z = 1, 2, \ldots, K)$, $(h = 1, 2, \ldots, G)$.where $\mathcal{B}$ is switch port identity that is represented in BLS table, and o for outlet (output) and I for inlet (input), as shown in Fig. 3(a). Clearly, SW, and S represents the switch element, and the stage, respectively, where n, K, and G denote the number of input or output switch's ports, the number of switches in a stage, and the number of stages in the network, respectively. Second, If there is ζ weight in one of the outlets of the switch element such $\mathcal{B}_i^o[SW_z, S_h] = \zeta$, as shown in Fig. 3(b). It means that some of the outlet has high traffic level. So, the control algorithm in the switch element will begin to reroute arriving packets that have address destined to the high traffic output port. Consequently, the traffic on the link will be reduce. The percentages of the packets' deflection γ of the total traffic arrivals λ at the switch element depend on the value that is determined by the user. Thereby, the traffic will pass $(1-\gamma)/n$ of the packet arrivals over each switch's inputs that has high weight outlet, and other γ/n of the packet arrivals will be distributed among links of other switch, where γ is deflection rate. Third, if the output port link is toward congestion such $\mathcal{B}_i^o[SW_z, S_h] = \Xi$ then the control process in the switch element has to make decision to add new weight value to its BLS input data $\mathcal{B}_j^I$.

Assumption 1:

In the same switch element, if any of output ports has Ξ weight value then all input ports will have ζ weight value.

All inputs of the switch element will have ζ weight, as shown in Fig. 3(c). The new weight value ζ for the input links in BLS table will transfer to other switches' output links in the stage S_{h-1}, such that:

$\mathcal{B}_j^I[SW_{z=i}, S_h] = \mathcal{B}_i^o[SW_{z=j}, S_{h-1}] = \zeta$, for $j = (1, 2, \ldots, n)$. Switch elements in the stage *h-1* will begin rerouting packets' arrival at switch's inputs. It means that the traffic will reduce on the switch element $[SW_{z=i}, S_h]$ in stage *h*.

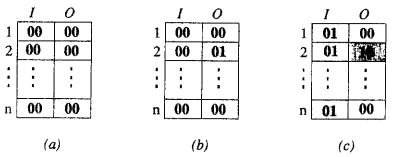

Fig. 3. Binary Link Status of the switch element

So, we should have to reduce congestion level on output port by controlling the packet arrival at whole input ports of the switch element . Therefore, we make partial restriction on packets arrival to reduce the high congestion level on output port. The packets arriving that has been deflected from its flows will distribute to other links. In Appendix A, the algorithm is to exchange the traffic weight between the outlets and inlets of the switch element.

The data share algorithm among BLS's tables in adjacent switch elements is described as follows.

In the interconnection between output ports at stage *h* and inputs ports at stage *h+1*, let the output port $\mathcal{L}_i^o$ $(i = 1, \ldots, n)$, be connected to the input port $\mathcal{L}_j^I$ $(j = 1, \ldots, n)$ for a switch element SW_z, z =(1, 2, . . ., n) in the next stage S_{h+1}, where h =(1, 2, . . ., K), n, z, and K are the total number of switch's output or input ports, the number of the switch in a stage and the number of the stages in the interconnection network, respectively. We have $\mathcal{B}_i^o[SW_z, S_h] = \mathcal{B}_j^I[SW_{z=i}, S_{h+1}]$ = [weight] because of the same point connection. The weight will transfer from output (outlet) in stage *h* to another input (inlet) in stage *h+1*. The switch elements will reconsider a new weight value that have been transferred to its outputs, and will take decision to reroute packets if the weight value was ζ.

3-2. Switch Routing Algorithm

The switch begins rerouting packets under heavy load condition when the link status has ζ weight which is represented by the BLS-table in switch's memory. This control procedure is active and associated with other switches only when one of the switch's outputs prone to congestion. The packet address has field addresses indicating the addresses for switch elements in interconnection network. Fig. 4 shows packet's header format passing through the interconnection network. Our concerning is the control scheme in which the packet's address can be changed in consequence of traffic congestion. This address field contains switch identification in which the packet will pass through interconnection network's stages, switch elements, and switch's output port identification. The usage of the other fields of the packet format are depending on the switch design.

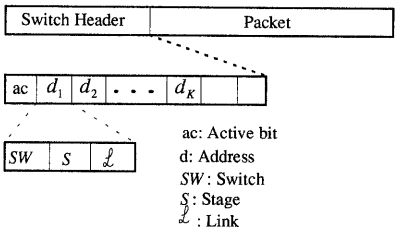

Fig. 4. Packet header format

Each address *d* has a set of three elements SW, S, and $\mathcal{L}$, such that $d_a = \{SW_z, S_h, \mathcal{L}_i^o\}$, where $a=(1, 2, \ldots, K)$. The switch's decision process for rerouting packets arrival can be made if the switch element has ζ weight which represents the traffic level of the output port in the switch's BLS table. The rerouting packet is achieved by distributing packets arrival, that have been deflected from high traffic switch, among switch's output ports. It means the packets arrival at the switch, that have address destined to the output port that has ζ value in BLS table, will be distributed among the output's switch ports, see Fig. 5.

Distributing packets can be made by changing the addresses $[d_1, d_2, \ldots, d_a]$ in the switch header. This change comprises the value of $\mathcal{L}$ that has ζ value in a stage *h* and SW in a stage *h+1* in a cyclic manner, see Fig. 5.

Assumption 2:

In the same switch element, If any of the output ports which has ζ weight value then the packets that have address destined to ζ weight's output will distribute among all links.

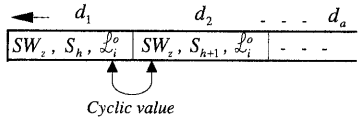

Fig. 5. Change address to rerouting packets arrival

The distribution packets arrival is done by switch element's control algorithm through changing packets address the output value $\mathcal{L}_i^o$ for i =(1, 2, . . ., n), and we have $|\mathcal{L}_i^o|_h = |SW_z|_{h+1}$. In Appendix B, the algorithm shown is to control packet arrivals that have addresses destined to high traffic link.

4. EXAMPLE

In this section, we consider an example to illustrate the proposed scheme. Fig. 6 shows four-stage switch architecture. We consider all switch elements are identical with four input/output ports. We suppose that one of the output ports of the switch 3 in the stage 3 (4x4) has a traffic

congestion level, e.g., output port 2. This output link is represented by thick solid line. The traffic congestion level on the output port will reflect on the traffic of the all input ports (Assumption 1). The dotted lines represent the input ports of the switch that has traffic congestion on one of its output port. The switch's control algorithm should limit packets arrivals among these ports (dotted lines) to reduce the traffic congestion on the switch element, and let the some packets to choose alternative route to distention.

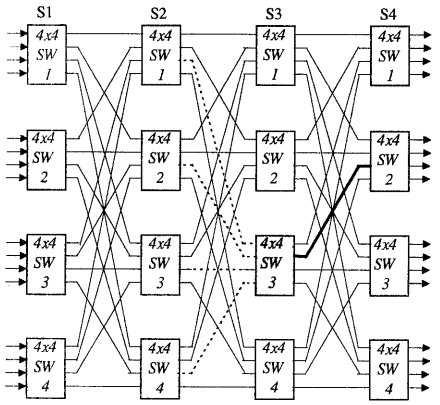

Fig. 6. Congestion in the switch architecture

The representation of the links' status is shown in Fig. 7. The monitoring traffic conditions are stored in the Binary Links Status table BLS which has mentioned in previous subsection. The BLS table of the switch SW_3S_3 has congestion level $\Xi(10)$ at the output link 2, this means the arriving packets exceed its throughput. This value will propagate to the other side of the BLS table of the switch (BLS's data for switch's input link) as ζ (01) value (Assumption 1). Therefore, the ζ value will transfer to other switch in previous stage (in this example, switch SW_1 *to* SW_4, link $\mathcal{L}_i^o$, and stage S_2). The control algorithm in each of the switch elements that the ζ value has been transferred to its BLS table will consider this value as heavy load output.

Therefore, the arriving packets in each switch elements of the stage 2 that have address for output port 2 will be distributed among the switch's output ports through rerouting the packets arrival. The arriving packets in stage 2 that have address for e.g. $d = \{SW_1, S_2, \mathcal{L}_2^o\}$, $\{SW_2, S_2, \mathcal{L}_2^o\}$, . . ., or $\{SW_4, S_2, \mathcal{L}_2^o\}$ have to be rerouted to choose other gates by modifying packet's header address. The new addresses of the arriving packets, e.g., in a switch SW_1 will take one of the following addresses $d_2 = \{SW_1, S_2, \mathcal{L}_1^o\}$, $\{SW_1, S_2, \mathcal{L}_2^o\}$, $\{SW_1, S_2, \mathcal{L}_4^o\}$, or $\{SW_1, S_2, \mathcal{L}_4^o\}$ and additional to the original addresses d_3, d_4.

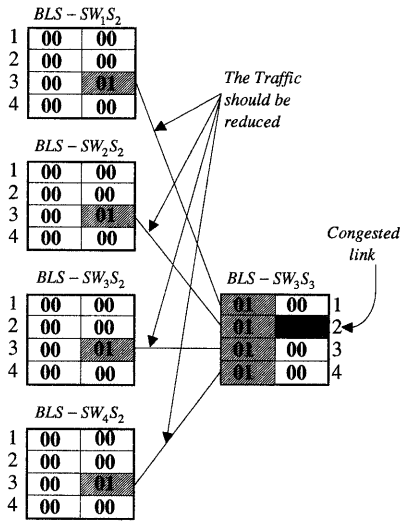

Fig. 7. Representation of the traffic status of the input/output ports

This kind of the adaptive rerouting algorithm relieves the interconnection network from consequence of the network congestion in nonuniform traffic patterns or in case of the switch element failures [5]. The rerouting bursty packets decreases the chance for congestion occurred in any of the switch elements for interconnection network. Since the congestion of the network is reduced, the network's throughput will be a significant increase.

5. PERFORMANCE ANALYSIS

In this section, we present queuing models to analyze the model performance in view of packet loss probability. We consider the switch element is an internally nonblocking switch. The input/output ports are identical with output queuing using a buffer of finite size H as Fig. 8. We consider the traffic at each input ports is nonuniform delivered to all output ports. Our target is to reduce packet loss in the switch under the condition that the switch has congestion in one of its output ports, and with small buffer size.

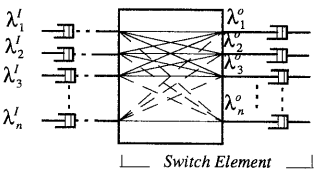

Fig. 8. Switch Element's model

5-1. Assumptions

We assume that the packet arrival at each input port of the switch element is random and independent. The arrival

process is assumed to be Poisson probability assignment. We analyze the effect of the packet loss probability in nonuniform traffic distributed to the output ports.

5-2. Nonuniform Traffic

The input loads are nonuniformly distributed over output's ports of the switch. The effect of a nonuniform traffic pattern has been investigated where many packets are destined for one destination such kind of traffic is called hot spot traffic [3] [14]. The packets will be discarded if the packets are exceed link capacity or they cannot be routed to the next stage.

In the following we analyze the impact of traffic under *Ants routing* strategy on the packet loss probability, for nonuniform traffic pattern. We assume that, the input traffic in each inputs is independent and Poissonian with same probability offered load of each input port ρ_j^I , $j=(1, 2, \ldots, n)$, where n is the total number of input ports of the switch element. We suppose that the traffic is nonuniformly distributed among output ports. The switch element has buffer in each input ports only. Assume that the queue is finite with a buffer capacity H of packets, and that the service discipline is first-in first-out (FIFO) policy. The packet arrival rate at each output λ_i^o , where $i=(1, 2, \ldots, n)$ is the superposition of input ports. The merging packets arrival stream from each of the sources is also Poisson process [15]. The packet loss probability can be estimated using an M/M/1/H model.

The probability of arrival packets from the input port j that are tended to choose specific output i is p_{ji} . Then the augmented matrices of the total rate arrivals distend to an output i is:

$$\begin{pmatrix} \lambda_1^o \\ \lambda_2^o \\ \lambda_3^o \\ \vdots \\ \lambda_i^o \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1i} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2i} \\ p_{31} & p_{32} & p_{33} & \cdots & p_{3i} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{j1} & p_{j2} & p_{j3} & \cdots & p_{ji} \end{pmatrix} \begin{pmatrix} \lambda_1^I \\ \lambda_2^I \\ \lambda_3^I \\ \vdots \\ \lambda_j^I \end{pmatrix} \quad (1)$$

and

$$\sum_{j=1}^{n} p_{ij} = 1, \quad \text{where } i=(1, 2, \ldots, n) \quad (2)$$

Also,

$$\lambda_i^o = \sum_{i=1}^{n} p_{ji}\lambda_j^I \quad \text{where } j=(1, 2, \ldots, n) \quad (3)$$

The steady-state packet loss probability PL for the output port i or of the buffer of the input in the next stage is given as:

$$PL_i = \frac{(1-\rho_i)\rho_i^H}{1-\rho_i^{H+1}} \quad (4)$$

where $\rho_i = \dfrac{\lambda_i^o}{\mu_i^o}$ (5)

The packet loss probability in a switch element sw is obtained as:

$$PL_{sw}[\text{packet loss}] = \frac{\sum_{i=1}^{n} PL_i\,\lambda_i^o}{\sum_{j=1}^{n}\lambda_j^I} \quad (6)$$

where $sw=(1, 2, \ldots, K)$

The total packet loss probability in a stage S is given as follow:

$$Pr_i[\text{packet loss in stage } S] = \frac{\sum_{sw=1}^{K}\sum_{i=1}^{n} PL_{sw}\,\lambda_i^o}{\sum_{sw=1}^{K}\sum_{j=1}^{n}\lambda_{j,sw}^I} \quad (7)$$

However, our policy is to reduce the packets arrival rate among the input port for the switch element that has high blocking level at any of the its output ports. We assume the deflection is uniform among the all input port of the switch element, and the arrival rate to input ports will be $\bar{\lambda}_{j,sw}^{I}$ where γ is deflection rate. From (2), the arriving packets to the congested switch element *(sw=C)* with deflection are:

$$\bar{\lambda}_{j,sw}^{I} = \begin{cases} \lambda_j^I \times (1-\gamma) & j=1, 2, \ldots, n \\ & sw = C \text{ (Congested switch)} \\ \lambda_j^I \times (n+\gamma-1)/(n-1) & j=1, 2, \ldots, N-n \\ & sw \neq C \end{cases} \quad (8)$$

The switch's output ports also reduce to:

$$\bar{\lambda}_{i,sw}^{o} = \begin{cases} \sum_{i=1}^{n} p_{ji}[\bar{\lambda}_{j,sw}^{I} \times (1-\gamma)] & \text{for } sw = c \\ \sum_{i=1}^{n} p_{ji} \times \bar{\lambda}_{j,sw}^{I}(n+\gamma-1)/(n-1) & \text{for } sw \neq c \end{cases} \quad (9)$$

From (4) we can obtain packet loss probability for one output port is:

$$PL_{i,\gamma} = \frac{(1-\rho_{i,\gamma})\rho_{i,\gamma}^{H}}{1-\rho_{i,\gamma}^{H+1}} \quad (10)$$

$$\rho_{i,\gamma} = \frac{\bar{\lambda}_{i,sw}^{o}}{\mu_{i,sw}^{o}} \quad (11)$$

In same way for deriving (6) and (7), we can compute the packet loss probability in the switch $PL_{sw,\gamma}$ that has congestion at one of its output ports under deflection of the input traffic load. The packet loss probability for one stage $PL_{i,\gamma}$ can be calculated with reduced the load on one of its switch elements.

5-3. Simulation Results

This subsection describes a simulation of the proposed algorithms, that has designed to compare the impact of packet loss probability with applying *Ants routing* strategy and without the *Ants routing* strategy. The results of packet loss probability that we have obtained are for one stage that has congestion in one of its switch's output port in nonuniform traffic pattern. Our target is to minimize the amount of the buffer required while maintaining a low cell loss probability, probably around 10^{-9}.

As an example we set the number of switch's ports n = 4, and the number of switches S in a stage h at $S = 4$, Fig. 9 shows the packet loss probability (without the *Ants routing*

scheme γ =0) as a function of load λ_n^o on one switch's ports for various values of buffer size and the offered load on each switch element in a stage is ρ = 0.9. The packet loss probability is aggravated by increasing in the load on one output port. It is required to increase buffer size to maintain low level of packet loss rate.

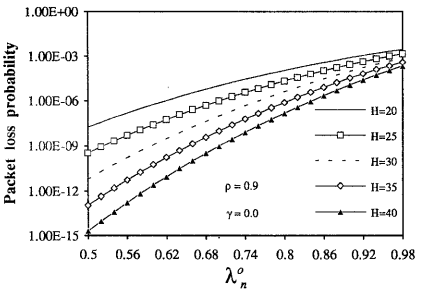

Fig. 9. Packet loss probability versus load on an output port

The buffer size increased is costly and out of our policy to minimize buffer size. Fig. 10 shows the relation between packet loss probability for specific stage and switch's ports n for various offered loads on the switch element (ρ =0.6, 0.7, 0.8, 0.9). We fix the value of H=10, λ_n^o =0.9, γ =0. From the Figure the packet loss decreases as n is increased. To keep the packet loss probability under 10^{-9} with small buffer size, the switch' ports (n) are required to be more than 8.

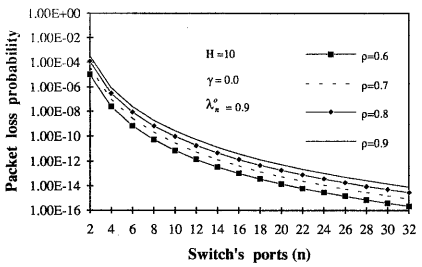

Fig. 10. Packet loss probability against the switch' ports (n), without *Ants routing* scheme

We now turn to the *Ants routing* scheme and how it can effect on packet loss probability. Fig. 11 shows the effect of buffer size for *Ants routing* scheme on packet loss probability and with various input rate deflections (γ =0, γ =0.1, γ =0.2, γ =0.3, γ =0.4). Note that γ =0 means the network without the *Ants routing* control mechanism. We fix offered load ρ on the switch element at 0.9, and the load λ_n^o destined for output port at 0.9. The result can lead to a large deflection in packet loss probability. Moreover, if we fix packet loss below 10^{-5}, we can reduce buffer size more than 40% of offered load on the switch.

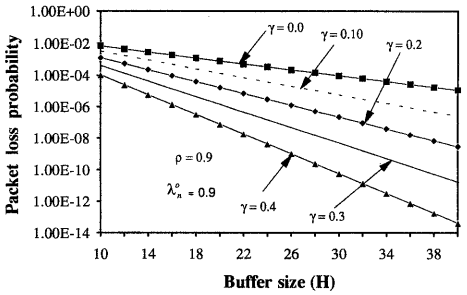

Fig. 11. Probability of packet loss versus buffer size (H) with *Ants routing* policy

For H=20 and λ_n^o =0.7, we can obtain significant improvement in the packet loss probability for a specific stage. Fig. 12 shows this improvement for various deflections in the input switch offered load (γ =0, γ =0.1, γ =0.2, γ =0.3, γ =0.4). To achieve lower packet loss rate of 10^{-9} under traffic load of ρ =0.9 is attained with γ =0.4.

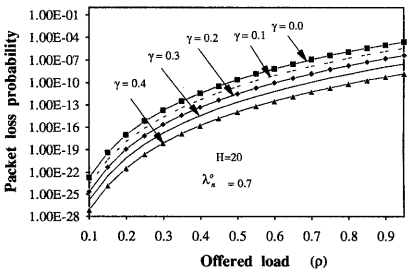

Fig. 12. Packet loss probability versus offered load (ρ)

Fig. 13 shows packet loss probability in a stage versus average packet arrival that is destined to the output link λ_n^o, where ρ = 0.9, H=30. The packet loss probability is small when the deflection rate γ is large. As the load at output link increase it is requited to increase γ to achieve packet loss of 10^{-9} .

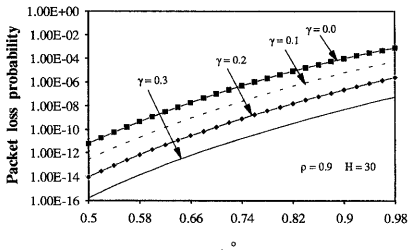

Fig. 13. Packet loss probability against the load on destined output port under Ants routing policy.

So far the analyses of performance of the switch element under *Ants routing* scheme were based on the input load deflection of the switch. The packets arrival at the switch that has congestion on one of its output will reduce to specific rate. This deflection rate will distribute to other switches' links that have low traffic congestion level.

6. CONCLUSION

We proposed a new self switch rerouting packet algorithms for congested network through source rate switch deflection called *Ants routing*. We have used (*BLS*) Binary Link Status table to represent the information about traffic status at switch's output ports. We asserted that the switch element is to be intelligent switch and to fortify switch control algorithms in consequence of traffic congestion. The switch network performance can be enhanced by adding a little information to the switch's software routing algorithms. Our technique can significantly reduce packet's loss probability and relieve network's congestion through rerouting packets to choose other paths. Moreover, we can use network resources more efficiently.

APPENDIX A

The algorithm to exchange the traffic weight between links in an internal switch element shows as follows.

TRAFFIC WEIGHT EXCHANGE BETWEEN LINKS

```
begin
  for i ← 1 to n do
    if ℬ_i^o[SW_z, S_h] = Ξ (10) ... (congestion level on one
      then                          of the switch's output)
        for j ← 1 to n do
          ℬ_j^I[SW_z, S_h] = ζ (01) ... (high traffic level – ζ weight will
        endfor                        transfer to the switch's inputs)
        for z ← 1 to n do (z = j)
            ℬ_i^o[SW_z, S_{h-1}] = ζ ... (ζ weight will transfer to the
        endfor                        switch's outputs of h – 1 stage)
    endif
  endfor
end
```

APPENDIX B

The distribution packets arrival is done by switch element's control algorithm through changing the output value. The control algorithm is shown as follow:

```
begin
  for i ← 1 to n do
    while ℬ_i^o = 01 do { Binary links status for switch element}
      for i ← 1 to n do
        d_1 = [SW_z, S_h, ℒ_i^o] {Changing output link's
                                  address in stage h}
        z ← i
        d_2 = [SW_z, S_{h+1}, ℒ_j^o] {Changing switch element
                                  address for stage h+1}
      endfor
    endwhile
  endfor
end
```

REFERENCES

[1] Satoru Ohta, and Ken-Ichi Sato, " Dynamic Bandwidth Control of the Virtual Path in an Asynchronous Transfer Mode Network", IEEE Trans. on Comm., vol. 40, no. 7, pp. 1239-1247, July 1992.

[2] Dimitri Bertsekas, Robert Gallager, "Data Network", Prentice-Hall, Inc., New Jersey 1992.

[3] Shigeo Urushidani, " Rerouting Network: A High-Performance self-routing switch for B-ISDN", IEEE J. Sel. Areas in Comm., vol 9, no.8, pp. 1194-1204, October 1991.

[4] A. Gersht and K. J. Lee, "A Congestion Control Framework for ATM Networks", IEEE J. on Selec. Areas in Comm., vol. 9, no. 7, pp. 1119-1130, Sep. 1991.

[5] Mark J. Karol, Salman Z. Shaikh, " A simple Adaptive Routing Scheme for Congestion Control in ShuffleNet Multihop lightwave networks", IEEE J. on Selec. Areas in Comm., vol. 9, no. 7, pp. 1040-1051, Sep. 1991.

[6] Tse-yun Feng, "A Survey of Interconnection Networks", IEEE Computer, vol. 14, pp. 12-27, Dec. 1981.

[7] C. Clos, "A Study of Non-Blocking Switching Networks", The Bell Sys. Tech. Jour., vol XXXII, pp. 406-424, March 1953.

[8] V. E. Benes, "On Rearrangeable Three-Stage Connecting Networks", The Bell Sys. Tech. Jour., vol XLI, no. 5, pp.1481-1492, Sep. 1962.

[9] Mark J. Karol, " Performance Analysis of a Growable Architecture for Broad-Band Packet (ATM) Switching", IEEE Trans. on Comm., vol. 40, no. 2, pp. 431-439, Feb. 1992.

[10] Kai Y. Eng , Mark J. Karol, Yu-Shuan Yeh, " A Growable Packet (ATM) Switch Architecture: Design Princibles and Applications", IEEE Trans. on Comm., vol. 40, no. 2, pp. 423-430, Feb. 1992.

[11] Lionel M. Ni, and Philip K. Mckinley, " A Survey of Wormhole Routing Techniques in Direct Networks", Computer, vol. 26, no. 2, pp. 62-76, Feb. 1993.

[12] Fouad A. Tobagi, Timothy Kwok, and Fabio M. Chiussi, "Architecture, Performance, and Implementation of the Tandem Banyan Fast Packet Switch" IEEE J. on Selec. Areas in Comm., vol. 9, no. 8, pp. 1173-1193, October 1991.

[13] G. Ramamurthy, and R. S. Dighe, " Distributed Source Control: A Network Access Control for Integrated Broadband Packet Networks", IEEE J. on Sele. Area in Comm., vol. 9, no. 7, pp. 990-1002, September 1991.

[14] P. Kermani and L. Kleinrock, " Virtual cut through: A new computer communication switching technique," Comput. Networks, vol. 3, pp. 267-286, 1979.

[15] Leonaed Kleinrock, "Queueing Systems Vol. I: Theory", John Wiley& Sons Inc., New York, 1975.