Adaptive Multipath Routing of Connectionless Traffic in an ATM Network

Josep Sole–Pareta * *Debapriya Sarkar* † *Jörg Liebeherr* † *Ian F. Akyildiz* ‡

* Departament d'Arquitectura de Computadors
Universitat Politecnica de Catalunya
Barcelona, Spain

† Computer Science Department
University of Virginia
Charlottesville, VA 22903

‡ School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332

Abstract

We address the problem of routing connectionless traffic in an ATM network and propose an adaptive multipath routing scheme that enhances the routing methods employed at the Inter-Working Units (IWUs). We present a scheme that distributes packets among multiple Virtual Paths (VPs) according to the utilization of the links of these VPs. The utilization of the VPs is determined by a feedback mechanism. We present simulation studies to show the effectiveness of the proposed adaptive multipath routing scheme.

1 Introduction

To support connectionless data traffic in B-ISDN one must consider the following three aspects:

Architecture. Three different schemes have been proposed for implementing connectionless data transport services on top of ATM Networks [14]:

(i) In the *On-Demand scheme*, a virtual channel connection (VCC) is established for each packet[1].

(ii) The *Semi-Permanent Connections scheme*, illustrated in Figure 1(a), assigns a Semi-Permanent Virtual Paths to each pair of Interworking Units (IWU). The result is a fully meshed network of virtual paths (VPs) that connect the IWUs, forming a virtual network of VPs on top of the ATM network.

(iii) The *Connectionless Virtual Overlay Network scheme*, shown in Figure 1(b), employs Connectionless Servers (CLS) that are connected by Semi-Permanent VPs. The CLSs perform routing functions for connectionless traffic.

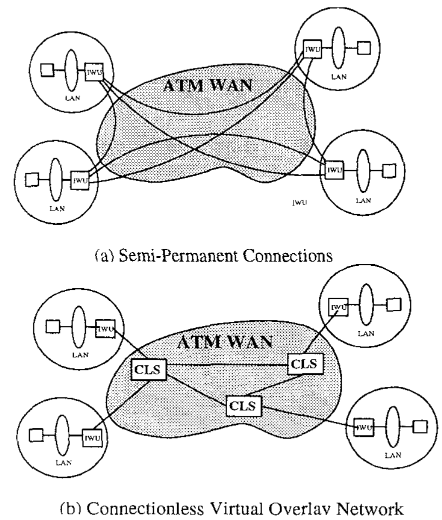


Figure 1: Approaches to LAN Interconnection.

Congestion Control Policy. A congestion control policy is required for efficient traffic management in each of the three architectures. The objective of this policy is to maximize the bandwidth allocation for connectionless traffic without violating the service guarantees given to connection-oriented traffic. Most approaches for bandwidth allocation fall into two categories: *bandwidth reservation* methods [1, 2, 13], where IWUs allocate a fixed amount of bandwidth for connectionless traffic, and *bandwidth regulation* methods [3, 4, 11, 17] where the bandwidth available to connectionless traffic is controlled by an adaptive feedback scheme that involves the ATM switches and the IWUs [3, 4, 17]. A recent study proposes a scheme

[0]The work of JSP was supported in part by a CIRIT (Generalitat de Catalunya) grant (expedient number EE92/2–338), and in part by the National Science Foundation under Grant No. INT–94033646. The work of IFA was supported in part by the National Science Foundation under Grant No. INT–94033646.

[1]We use the term **packet** to refer to a *Connectionless Protocol Data Unit* or CL_PDU.

that combines bandwidth reservation and regulation [6].

Routing. In a virtual networks of VPs with VP endpoints as nodes and VPs as links, two routing methods are considered for improving a connectionless ATM service: With *alternate path routing* traffic is sent primarily on a single path, but alternate paths can be selected if the primary path is congested. With *multipath routing* multiple paths are used simultaneously. Note that multipath schemes can be implemented on a per-connection basis or on a per-packet basis. Recently it was shown that per-packet multipath routing [9] yields a better performance, even though the method incurs resequencing delays.

We study a per-packet multipath routing scheme which couples congestion control with the routing. The first scheme that considered a coupling of congestion control and routing is proposed in [16]. The scheme provides each pair of IWUs with several parallel VPs, initially without any bandwidth allocation. When a burst must be sent, a broadcast request process is initiated on all paths. All paths independently try to allocate the bandwidth requested. The packet is transmitted along one of the VPs that have sufficient bandwidth available. The rest of the VPs release their bandwidth. However, bandwidth reservation on a burst-by-burst basis incurs control overhead and may result in excessively long delays if the propagation delay between the source-destination pair is long.

We address the problem of connectionless traffic management in ATM networks CLS and the semi–permanent VP architectures. We propose an adaptive multipath routing mechanism that achieves efficient statistical multiplexing of connectionless traffic and incorporates congestion control based on a periodic feedback mechanism. Packets are distributed among multiple VPs according to the information obtained from the feedback scheme.

The remaining parts of the paper are organized as follows. In Section 2 we introduce the adaptive multipath routing scheme. In Section 3 we study the performance of the new method through simulation experiments. In Section 4 we conclude the paper.

2 Adaptive Multipath Routing of Connectionless Traffic

We consider an architecture with Semi-Permanent Connections where each IWU can choose between different VPs for transmitting a packet. Cells of the same packet are transmitted on the same VP. The steps of our method are outlined as follows:

Step 1: Multiple VP Set-Up. We assign several semi–permanent VPs, placed on different routes, for each IWU–IWU pair that interconnects connectionless sources.

Step 2: Bandwidth Reservation for Connectionless Traffic. To prevent connection–oriented traffic from affecting the bandwidth available to connectionless traffic, a fixed amount of bandwidth is reserved for exclusive use by connectionless traffic. In our case, this amount of bandwidth is split among the semi-permanent VPs assigned to the source IWU-destination IWU pairs. The bandwidth reservation should be pessimistic to account for the burstiness of connectionless traffic. We propose to use the estimation method presented in [15].

Step 3: Adaptive Multipath Routing. At a source IWU, the arriving packet is probabilistically routed on any one of the multiple routes. The probability for choosing a particular VP depends on the relative utilization of the links on the route of the VP. The probabilities for the VPs are dynamically updated using the feedback mechanism suggested in [4].

The multipath routing algorithm is executed in the AAL layer of the IWUs. This means that each IWU maintains a table with all possible destinations, each associated with a VPI, in this layer. This table is used to set the VPI field of the cells into which a packet is segmented. In using the adaptive multipath routing an IWU takes the following actions:

- The IWU updates the table that associates each VP with the corresponding destination every time a packet is completely delivered to the ATM layer.
- The probabilities of the VPs are re-calculated periodically by feedback cells generated by the destination IWU. A feedback cell traverses the route of a VP in reverse order, collects information on the load the links on a VP, and makes the information available to the source IWU.

We can assume that these actions will take place after the segmentation, and before setting the VPIs of the cells of a new packet. The process of updating the VPI consists of simply checking the SAR Type (Segmentation And Reassembly Type) of the SAR_PDU header if AAL 3/4 is being used, or the PTI (Payload Type Indicator) of the ATM cell header if AAL 5 is being used [8], and running the procedure when the last cell of a packet is delivered to the ATM layer, which will give us the new VPI.

Figure 2 shows a flow-chart describing the process of updating the VPI for the general case of n VPs per IWU–IWU pair. Assuming that the current VPI value in the routing table is α for a given destination D, as a result of running the algorithm, the new VPI value may remain α or change to β according to the current values of the n probabilities $p(i)$ and the current value of a random variable r, which is uniformly distributed between 0 and 1.

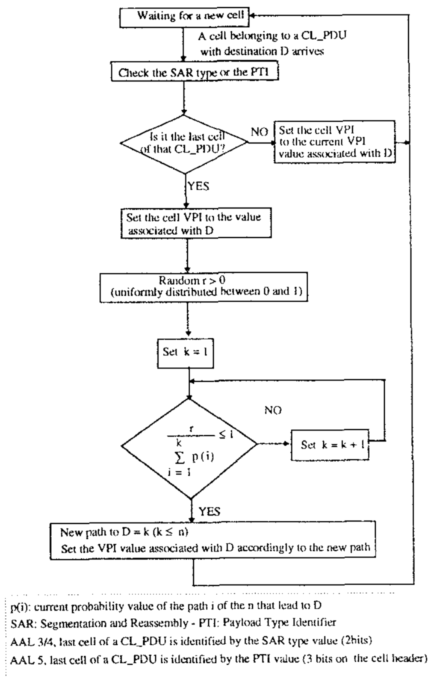


Figure 2: Adaptive Multipath Routing Algorithm.

Feedback Mechanism. For updating the routing probabilities we require that the IWUs receive congestion information from the network. Our metric for network congestion is the one suggested in [4], i.e., the maximum buffer occupancy at a switch encountered by a packet on its route. We propose a feedback procedure where control cell that indicate the congestion of a route are sent by the destination at periodic intervals. The payload of the control cells contains the maximum average utilization of the buffers at any switch on the route of a VP, where the averaging interval is given by the time interval between two consecutive control cell arrivals. When receiving a control cell, a switch compares its buffer utilization with the current contents of the control cell, and writes the greater value into the control cell.

3 Performance Evaluation

We have investigated the efficiency of the multipath scheme in a set of simulation experiments. We have implemented a simulator to model the ATM network shown in Figure 3.

The simulated network has a single *source IWU*, a single *destination IWU* and n intermediate ATM switches. The IWUs and ATM switches are connected by $2n$ links with a capacity of 150 Mb/s each. The buffer size of each AT switch is set to 50 cells. The

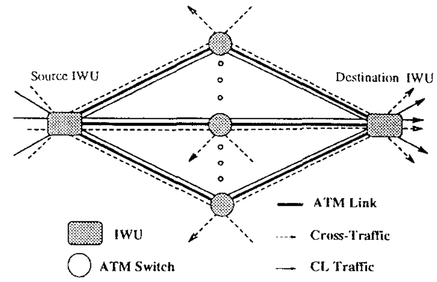


Figure 3: The ATM Network Model (Three–Paths Scenario).

source and the destination IWUs are linked by n parallel but non-overlapping Virtual Paths (VPs). In the simulations, we consider a two–paths scenario ($n = 2$) and a three–paths scenario ($n = 3$). The paths are indicated as VP1, VP2, and VP3. The two links of a VP are indicated as L_{i1} and L_{i2} for VPi (i$= 1, 2, 3$).

We distinguish two types of network traffic:

- *CL Source Traffic*: There is a fixed number of connectionless sources which generate packets with exponentially distributed size with mean N_{CL}. At the source switch, a packet is fragmented into cells and allowed into the network at (fixed) rate $R_{CL_{peak}}$. Packet arrivals are Poisson with rate $R_{CL_{ave}}$. The numerical values of the parameters are summarized in Table 1.
- *Cross Traffic*: The cross traffic on each link is statistically independent and represented by a two–state model (HIGH state, LOW state). In the HIGH (LOW) state, cells are generated with exponential interarrival times with rate $R_{X_{high}}$ ($R_{X_{low}}$). The duration of the HIGH (LOW) state is exponentially distributed with mean $T_{X_{high}}$ ($T_{X_{low}}$). We distinguish 'congested' and 'uncongested' links with different parameters for the cross traffic. The parameters for the links of the same VP are assumed to be identical. The paramters are given in Table 2.

Next we discuss the results of two simulation experiments. In the first experiment, we compare the performance of the *single path option* against the *multipath option*. Our performance measure is the cell loss ratio (CLR) of the connectionless sources, that is, the proportion of connectionless cells that are dropped in the network. In the second experiment, we study the stability of the adaptive routing method.

3.1 Experiment I

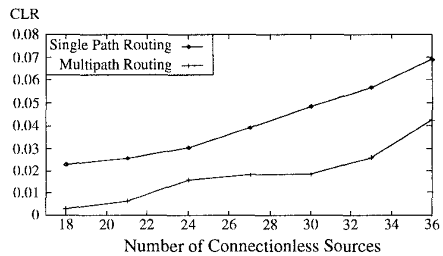
In this experiment, we measure the sensitivity of the cell loss ratios to changes of the number of connectionless sources. We assume the links of VP1, L_{11} and L_{12}, are congested while the other links are uncongested. The number of connectionless sources is varied between $S_{CL} = 18$ and $S_{CL} = 36$ in the two–paths scenario, and between $S_{CL} = 18$ and $S_{CL} = 66$

$R_{CL_{peak}}$	10 Mb/s
$R_{CL_{ave}}$	50 Kb/s
N_{CL}	100 cells

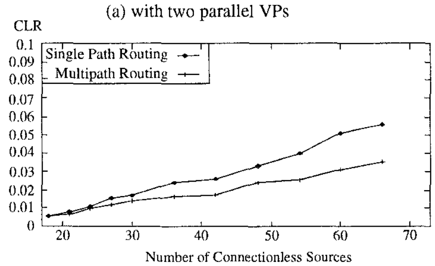
Table 1: Parameters of Connectionless Sources.

	Uncongested Link	Congested Link
$R_{X_{high}}$	40 Mb/s	120 Mb/s
$R_{X_{low}}$	30 Mb/s	30 Mb/s
$T_{X_{high}}$	1.06 seconds	1.41 seconds
$T_{X_{low}}$	0.35 seconds	1.41 seconds

Table 2: Cross Traffic Parameters.



(a) with two parallel VPs



(b) with three parallel VPs

Figure 4: Effect of Varying Source Traffic.

in the three–paths scenario. The CLR values obtained in the experiment are shown in Figure 4. We observe that the multipath option consistently performs better, i.e., has a lower CLR value than the single path option. In the three–paths scenario, both options have similar CLR values initially, but with the increase in the number of connectionless sources, the multipath option is clearly seen to be superior.

3.2 Experiment II

In this experiment, we study the transient behavior of the multipath option to demonstrate that the algorithm is stable. We plot the probabilities for selecting any of the three VPs as calculated by the feedback mechanism over a period of time. The setup for this experiment is different from the previous experiments. We only consider the three-path case and assume that all VPs have identical cross traffic rates of $R_{X_{high}} = 120Mb/s$ and $R_{X_{low}} = 30Mb/s$. Also, the state transitions of the cross traffic on the links are now deterministic. Initially, all links are in the LOW state. In the experiment, all links of a particular VP undergo state transitions simultaneously. The order of transitions is indicated in Figure 5. The parameters for the connectionless sources are as given in Table 1. The number of connectionless sources is fixed at $S_{CL} = 42$. The simulation results are shown in Figures 5 and 6.

We now discuss the experiment in greater detail:

- t = *0.00 sec*: The cross traffic at all links start in the LOW state. As the traffic load on each of the VPs is the same, the feedback mechanism sets the routing probability to 1/3 on the average.

- t = *0.10 sec*: The cross traffic at the links of VP2, L_{21} and L_{22} goes to the HIGH state. This raises the traffic load on VP2, leading to congestion. The feedback mechanism detects the congestion and causes the routing probability of VP2 to be at 0 for most of the interval, and those of VP1 and VP3 to be at 0.5. Due to the random arrival of cells at the switches and partly due to the adaptive nature of the multipath scheme, it can be observed that the routing probabilities fluctuate.

- t = *0.15 sec*: The cross traffic at links L_{31} and L_{32} of VP3 changes from LOW to HIGH. Now, the links of both VP2 and VP3 have high traffic levels compared to the links of VP1. The buildup of congestion causes the routing probabilities of VP2 and VP3 to be at 0.17 during most of the interval, while that of VP1 is at 0.66.

- t = *0.20 sec*: The cross traffic at the links of both VP1 and VP2 transition to the HIGH state. Now, links are in the HIGH state and are equally likely to suffer from congestion. Therefore their routing probabilities of the VPs are at 0.33 most of the time.

- t = *0.25 sec*: The cross traffic at the links of VP1 and VP2 return to the LOW state. This situation is similar to that at $0.10 \leq t \leq 0.15$ *sec*. The routing probabilities of VP1 and VP2 fluctuate between 0.5 and 0.33, while that of VP3 fluctuates between 0 and 0.33.

- t = *0.30 sec*: The cross traffic at links L_{31} and L_{32} of VP3 returns to the LOW state. All links in the network are now uncongested. Accordingly, the routing probabilities of all VPs are now at 0.33 on the average.

The short term dynamic behavior of the feedback mechanism is shown in Figure 6, which plots the routing probabilities over the time period $0.1secs \leq t \leq 0.13$ *secs*. In the figure, the cross traffic experienced by VP1 and VP3 are in the LOW state and that at VP2 is in the HIGH state. We observe that the routing probabilities of the three VPs alternate between two different sets of values.

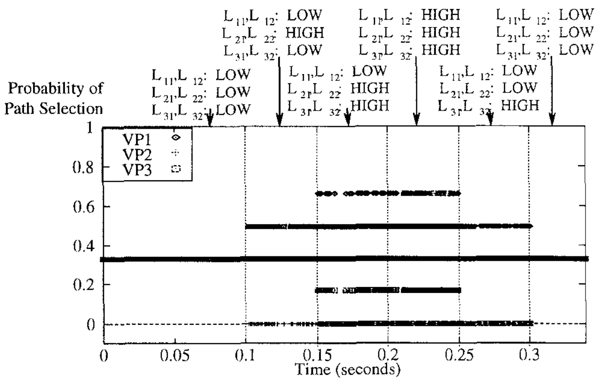


Figure 5: Transient Behavior of the Probabilities.

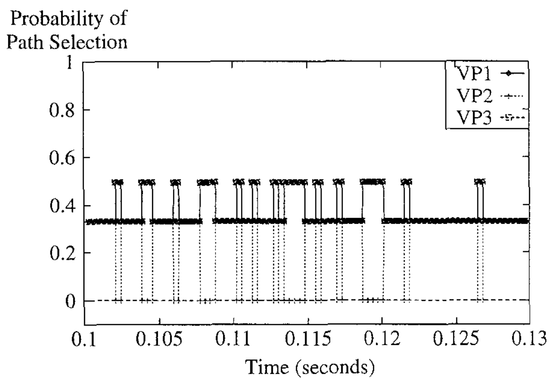


Figure 6: Transient Behavior of the Probabilities (Enlargement from Figure 6 for the Time Interval [$0.1 \leq t \leq 0.13$]).

4 Conclusions

Our proposal to support connectionless traffic in an ATM network is based on the approach of defining a virtual network of VPs overlaid on top of an ATM network. The management of the virtual network is performed by the IWUs. Presently, the virtual network is defined in such a way that every connectionless source IWU-destination IWU pair is interconnected by at least two independent paths. A fixed amount of bandwidth is allocated among the VPs connecting the same source IWU-destination IWU pair. We have proposed a scheme that detects congestion in the VPs and adapts the traffic distribution between the multiple VPs accordingly. We have performed simulations to show that our adaptive multi-path routing scheme provides a lower cell loss ratio for connectionless traffic in the face of congestion as compared to static multipath routing.

References

[1.] G. Bianchi, F. Borgonovo and V. Trecordi. "An Optimal Bandwidth Allocation Algorithm for Remote Bridging of FDDI Networks Across B_ISDN". *Proc. IEEE Globecom'92*, pages 1623–1627, 1992.

[2.] P.E. Boyer and D.P. Tranchier. "A Reservation Principle with Applications to the ATM Traffic Control". *Computer Networks and ISDN Systems*, Vol. 24, 1992, pages 321–334.

[3.] M. Gerla, T.-Y. C. Tai and G. Gallassi. "Internetting LANs and MANs to B_ISDN for Connectionless Traffic Support". *IEEE Journal of Selected Areas in Communications*, Vol. 11, No. 8, pages 1145–1159, August 1993.

[4.] Y. Gong and I. F. Akyildiz. "Dynamic ATM Traffic Control Using Feedback and Traffic Prediction". *Proc. IEEE Infocom '94*, pages 91–99, June 1994.

[5.] R.-H. Hwang, J. F. Kurose, and D. Towsley. "MDP routing in ATM Networks Using Virtual Path Concept". *Proc. IEEE Infocom '94*, pages 1509–1517, June 1994.

[6.] C. Ikeda, and H. Suzuki. "Adaptive Congestion Control Schemes for ATM–LANs". *Proc. IEEE Infocom '94*, pages 829–838, June 1994.

[7.] ITU–T/CCITT SG. XVIII, "Draft Recommendation I.364: Support of Broadband Connectionless Data Service on B–ISDN". Geneve, March 1993.

[8.] CCITT SG. XVIII, "Draft Recommendation I.363: B–ISDN Adaptation Layer (AAL) Specification". Geneve, March 1993.

[9.] R. Krishnan and J.A. Silvester. "Choice and Allocation Granularity in Multipath Source Routing Schemes". *Proc. IEEE Infocom '93*, pages 322–329, June 1993.

[10.] R. Krishnan and J.A. Silvester. "Resource Allocation in Broadband Networks – Cell, Burst or Connection Level?". *Proc. IEEE ICC'94*, pages 86–90, 1994.

[11.] H.T. Kung. "Use of Link by Link Flow Control in Maximizing ATM Network Performance: Simulation Results". *Proc. of the IEEE Hot Interconnects Symposium'93*, August 1993.

[12.] T. Van Landegem and R. Peschi. "Managing a Connectionless Virtual Overlay Network on Top of ATM Network". *Proc. IEEE ICC'91*, Paper No. 31.5, 1991.

[13.] L. Mongiovi, M. Farrell and V. Trecordi. " A Proposal for Interconnecting FDDI Networks through B_ISDN". *Proc. IEEE Infocom'91*, pages 1160–1167, April 1991.

[14.] P. Newman. "ATM Local Area Networks". *IEEE Communication Magazine*, pages 86-98, March 1994.

[15.] J. Sole-Pareta. "Connectionless Traffic on Top of an ATM Network". *Technical Report*, RR-93/30 Departament d'Arquitectura de Computadors, December 1993.

[16.] H. Suzuki and F.A. Tobagi. "Fast Bandwidth Reservation Scheme with Multi-Link and Multi-Path Routing in ATM Networks". *Proc. IEEE Infocom'92*, pages 2233–2240, May 1992.

[17.] N. Yin and M. G. Hluchyj. "On Closed-Loop Rate Control for ATM Cell Relay Networks". *Proc. IEEE Infocom'94*, pages 99–108, June 1994.