

August 2000

An examination of IP/ATM cut-through forwarding in dynamically routed networks

P. Boustead

University of Wollongong, boustead@uow.edu.au

Joe F. Chicharo

University of Wollongong, chicharo@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/infopapers>



Part of the [Physical Sciences and Mathematics Commons](#)

Recommended Citation

Boustead, P. and Chicharo, Joe F.: An examination of IP/ATM cut-through forwarding in dynamically routed networks 2000.

<https://ro.uow.edu.au/infopapers/206>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

An examination of IP/ATM cut-through forwarding in dynamically routed networks

Abstract

Multiprotocol over ATM (MPOA), IP switching and multiprotocol label switching (MPLS) have distinctly different mechanisms for cut-through packet forwarding. MPOA and IP switching use flow-based cut-through (FBC) forwarding while MPLS uses routing table linked cut-through forwarding (TLC). This paper examines the sensitivity of each these cut-through forwarding mechanisms to changes in underlying routing tables. We examine a scenario where a congestion-sensitive dynamic routing protocol, such as OSPF optimised multipath, leads to frequently changing routing tables. We show that FBC forwarding reacts significantly worse than flow length distributions predict, taking up to 1200 seconds to react to route changes and forward at least 50% of packets on the new route. Flow characteristics are examined to determine ways to improve FBC sensitivity. We show that implementing a maximum flow length of 200 seconds improves the response to route changes significantly with a minimal decrease in the number of switched packets (1.5%).

Keywords

asynchronous transfer mode, optimisation, packet switching, protocols, telecommunication congestion control, telecommunication network routing, telecommunication traffic

Disciplines

Physical Sciences and Mathematics

Publication Details

This paper originally appeared as: Boustead, P & Chicharo, J, An examination of IP/ATM cut-through forwarding in dynamically routed networks, International Conference on Communication Technology Proceedings, WCC - ICCT 2000, 21-25 August 2000, vol 1, 834-840. Copyright IEEE 2000.

An Examination of IP/ATM Cut-through Forwarding in Dynamically Routed Networks

Paul Boustead, Joe Chicharo
University of Wollongong
Northfields avenue, Wollongong, NSW, Australia
Email: paul@snrc.uow.edu.au

Abstract

Multiprotocol over ATM (MPOA), IP Switching and Multiprotocol Label Switching (MPLS) have distinctly different mechanisms for cut-through packet forwarding. MPOA and IP Switching use flow based cut-through (FBC) forwarding while MPLS uses routing table linked cut-through forwarding (TLC). This paper examines the sensitivity of each these cut-through forwarding mechanisms to changes in underlying routing tables. We examine a scenario where a congestion sensitive dynamic routing protocol, such as OSPF Optimised Multi-Path, leads to frequently changing routing tables. We show that FBC forwarding reacts significantly worse than flow length distributions predict, taking up to 1200 seconds to react to route changes and forward at least 50% of packets on the new route. Flow characteristics are examined to determine ways to improve FBC sensitivity. We show that implementing a maximum flow length of 200 seconds improves the response to route changes significantly with a minimal decrease in the number of switched packets (1.5%).

Keywords

Label Switching, MPOA, MPLS, Dynamic routing.

1. Introduction

MPLS [1], IP Switching, MPOA and several label switching protocols are all designed to provide scalable IP packet forwarding. However, they differ substantially in the way in which they route and forward packets. This paper examines the link between forwarding and routing decisions in these protocols. Forwarding is not always explicitly linked directly to routing decisions. This paper investigates IP label switching protocols used in conjunction with adaptive routing protocols that react to congestion or other network conditions and lead to dynamically changing routing tables.

Whereas traditional IP forwarding requires a longest-matching-prefix lookup to determine the next hop for a

packet, label swapping forwarding uses a direct lookup mechanism similar to that used by ATM. To enable label swapping forwarding an additional field must be carried with the packet, this field is termed a label or a tag. This label is added to a packet when the packet first enters a label-swapping network. In some cases the labels are placed in the VPI/VCI fields in an underlying ATM network. Switches within the network use the labels attached to packets and the incoming port number to lookup a table to determine the output port and a new label. The label field in the packet is over-written (or swapped) with the new label. Label swapping protocols do not forward all packets using labels in all switches. At some point network layer forwarding, which is also called routing or IP level forwarding, is required. The path followed by packets that use label swapping forwarding is also commonly called a cut-through path or a label switched path.

We group label switching protocols into two categories: Flow Based Cut-through (FBC) forwarding, and Table-Linked Cut-through (TLC) forwarding. FBC forwarding makes the routing decision when the cut-through flow is created and subsequent packets in the same flow will follow the same path. FBC protocols include IP Switching [2] and MPOA. TLC label switched paths are linked directly to routing protocol entries. These label switched paths are created and modified upon receipt of routing protocol messages. TLC capable protocols include MPLS, IP Navigator [3] and Tag Switching [4]. These two groups of protocols will perform differently in situations where the underlying routing protocol changes frequently.

Current IP routing protocols are insensitive to changes in network conditions such as congestion, and only provide equal cost routes to destinations based on fixed weightings defined at each hop. There is work underway, by the OSPF working group in the IETF, applying dynamic routing for best effort services to allow routing of datagrams around heavily loaded links [5]. The introduction of such routing protocols will produce a situation where the lowest cost route to a particular destination may change frequently depending on network load. If a route table entry changes due to changing network congestion it may not be beneficial to reset all label switched paths,

but just forward new label switched paths over the newly calculated route.

In this paper we examine FBC and TLC forwarding mechanisms with backbone traffic traces. The time response to a single route change in the underlying routing protocol is examined first. We define the time response as the time period before 50% of packets are forwarded on the new route. The characteristics of the traces are then examined to explain the poor performance of FBC forwarding. Finally we examine the effect of changing FBC flow detection parameters and the inclusion of a maximum flow length to improve the time response while minimising the number of cells forwarded at the IP layer. TLC forwarding will react immediately to routing table changes due to the nature of the forwarding mechanism. FBC forwarding will react significantly slower since existing flows will follow the old route. However, we find that FBC performs even worse than expected from published flow length distributions [6]. In our worst case example FBC forwarding took over 1200 seconds to start forwarding 50% of the packets on the new route after a route change. We find that implementing a maximum flow length does not result in a significant reduction in the number of switched cells while dramatically improving the response time.

The remainder of this paper is laid out as follows. Section 2 describes FBC forwarding and TLC forwarding. Section 3 discusses current advances in dynamic routing protocols. The trace driven experiments are described in Section 4. Results examining the performance of FBC forwarding are presented in Section 5. Section 6 concludes.

2. Cut-through Forwarding Classification

There are many different cut-through forwarding mechanisms in current label switching protocols. These mechanisms can be classified into FBC and TLC forwarding. FBC links cut-throughs to flow definitions; whereas, TLC links cut-throughs to routing information. This section examines these mechanisms and provides examples.

2.1 Flow Based Cut-through Forwarding

Flow based cut-through forwarding (FBC) abstracts the forwarding mechanism from routing table entries. When a cut-through flow is created the routing decision is made for that flow, a fixed label switched cut-through path is created and all subsequent cells in that flow follow this fixed route. In this section we briefly examine two FBC protocols, IP Switching and MPOA. While MPOA is not usually termed a label switching protocol we include it here because it allows creation of ATM cut-through VCs, which are similar to label switched paths.

MPOA [7] uses data driven techniques to determine if a cut-through flow should be created. The IP address of the source and destination MPOA Client (MPC) is used to de-

fine flows. The decision to create a cut-through between MPCs is based on a packet threshold parameter. The first packets in a flow, before the threshold is exceeded, are forwarded hop-by-hop by IP routers between Emulated LANs to the destination. The ingress node of the MPOA network determines if an end-to-end cut-through is necessary. If the number of datagrams in the flow exceeds a threshold value within a certain time period the ATM address of the destination MPC is determined and a cut-through is created. All subsequent packets follow this cut-through.

IP Switching also uses a packet threshold to determine if a cut-through path should be created. When the threshold is exceeded a cut-through path is created between itself and an adjacent IP Switch. When this happens in all IP Switches along the path an end to end cut-through is created. As with MPOA, once the cut-through VC is created all subsequent packets on that flow follow this fixed route independent of underlying routing table changes. Flows time out and are removed if no packets are received in a certain time period (usually 60 seconds).

2.2 Table-linked Cut-through Forwarding

Table linked cut-through forwarding links cell level cut-through paths directly to routing table information. Examples of protocols which support TLC forwarding are MPLS, Tag Switching and IP Navigator. For simplicity we only describe the ATM based version of the Tag switching forwarding mechanism which demonstrates the concept of TLC forwarding.

The trigger for the creation of a Tag Switching cut-through is either the receipt of a standard IP routing protocol packet advertising a new route, or a proprietary tag distribution protocol packet [8]. The main component of a tag-switching network is a Tag Switch. A Tag Switch maintains a Forwarding Information Base (FIB), and a Tag Information Base (TIB) as shown in Figure 1. The FIB is populated using information from routing protocol messages, and is similar to the routing table in a standard IP router. The TIB is essentially the switch VC table. Tag Switches bind all entries in the FIB with tags in the TIB. The first hop Tag Switch performs network layer forwarding to find the correct entry in the FIB. The associated tag in the TIB will then be placed in the cells VPI field, and the datagram is then forwarded through the ATM switch using this tag. Subsequent Tag Switches will have pre-setup bindings between this tag and a tag for its next hop router to the destination. Thus, ATM will switch the datagram to its destination. Binding of tags to entries in the FIB does not eliminate the need for network layer forwarding [4]. Traditional IP forwarding must be performed at the entry into a Tag Switching network to untagged datagrams. Additional IP forwarding is required since routing tables do not necessarily maintain a route to all possible destinations due to aggregation within IP

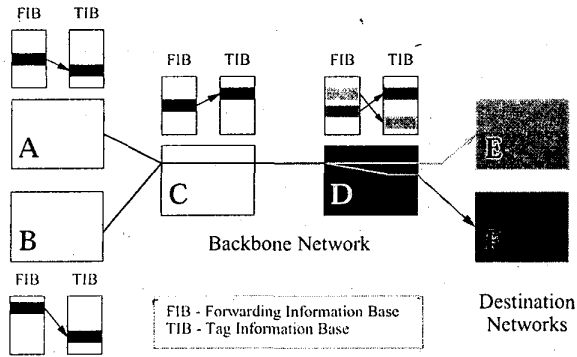


Fig. 1. Operation of Tag Switching

routing tables.

3. Dynamic Routing

Popular IP routing protocols such as OSPF and BGP are insensitive to changes in network conditions such as congestion, and are only capable of providing equal cost routes to destinations based on fixed weightings defined at each hop. Previously there have been attempts at adding delay based dynamic routing to the ARPANETs SPF routing protocol [9]. These approaches were shown to provide effective routing based congestion control and substantial improvements in delay and throughput. However, they suffered problems due to coarse granularity of routing changes leading to severe oscillations [5]. Currently there is work in progress developing a utilisation driven dynamic routing protocol with finer granularity.

An IETF working group is working on a proposal to allow best effort traffic to be dynamically routed around congested links. This proposal called OSPF Optimized Multi-Path (OSPF-OMP) [5] works by propagating information about load levels in links across the Autonomous System. The flooding of this information is facilitated by use of OSPF's Opaque Link State Advertisement option (Opaque LSA) [10]. The link state information includes: link load, packet loss, and link capacity. Using this information each router can use the internal link state topology map to determine the lowest-cost route to the destination according to current network conditions and the utilisation of links within the network. A finer granularity of control is facilitated by the ability to forward a proportion of packets on multiple links.

Use of dynamic routing leads to a situation where the best or lowest cost route in terms of delay, utilisation or loss will change frequently as network conditions change. Packet-based forwarding will take advantage of these changes in network conditions and forward packets on the appropriate route. However, the FBC approach will react more slowly to these changes. It would be impractical to tear down all FBC label switched paths ev-

ery time the underlying routing protocol changed due to changing network congestion. This paper examines the case where existing label switched paths are maintained until they timeout. Only packets belonging to new label switched paths will follow the new route.

4. Methodology

In order to examine the sensitivity of FBC forwarding mechanisms we used backbone traffic traces along with flow detection code. Of most interest was the response of each forwarding type to changes in the underlying routing table. TLC forwarding will obviously perform significantly better than FBC. The purpose of these tests is to quantify how much longer it takes FBC forwarding to adapt to routing table changes, and to identify possible ways to reduce this sensitivity without significantly increasing the number of packets requiring network layer forwarding.

Flow based studies in literature such as [6] analyse flow statistics of traffic traces. This information can be used to give a rough expected result. In [6] multiple traces are analysed using host pair flows. This study shows 50% of flows are shorter than 7 seconds in duration and 95% of flows are shorter than 1000 seconds (trace: SD-int PM in [6]). If all flows have the same traffic rate then FBC forwarding could be expected to react quickly with over 50% of packets transferring to the new route within 10 seconds of the route change.

We used traffic traces obtained from the National Laboratory for Applied Network Research (NLNR) over a 3 year period. The dates, duration and size of these traces is shown in Table I. Each of these traces was generated at the western USA Federal Internet Exchange (FIX-West) backbone node. Figure 2 shows how the traces were processed. We collected trace characteristics of the whole trace, then passed it through a FBC or TLC forwarding mechanism which forwards the packet on route 1 or route 2 depending upon the routing table. The "routing table" is set to forward all packets on route 1 until a route change trigger is received after which the routing table is changed to forward packets on route 2. Since FBC forwarding is not directly linked to routing tables then some packets will still follow link 1 after the route change.

The FBC forwarding mechanism uses a packet threshold for flow detection which was set to 10 packets in 60 seconds, these values are commonly used in literature [2]. A flow timeout is used to end flows due to inactivity. We used a flow timeout of 60 seconds for the initial examination, however, we varied this in later tests to examine the effect. A maximum flow length was also implemented whereby flows are ended if the flow-time exceeds x seconds even if they are still active. This facility was investigated in the last experiment described in the results section.

TABLE I
TRAFFIC TRACES USED

Date of Trace	Duration	Number of packets
21/6/1995	1625	132202033
28/2/1996	770	79583186
9/1/1997	1084	7302564
20/11/1997	1154	110409204

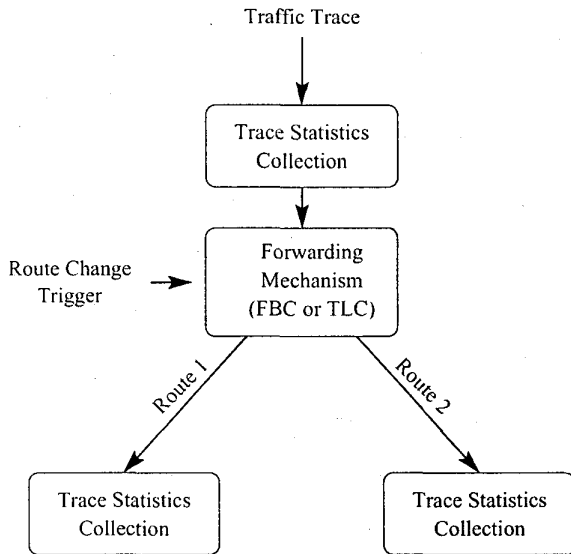


Fig. 2. Traffic trace analysis

5. Traffic Analysis

This section presents the results of our examination of FBC forwarding. Initially the time response due to a single route change is examined. The trace characteristics are then examined to explain the results for FBC forwarding. Finally flow detection parameters are altered to determine the effect on the results observed.

The time response to a single route change is shown in Figure 3 for the four different backbone traffic traces. All packets are forwarded on route 1 (see Figure 2) for the first 100 seconds. After 100 seconds the routing table is changed to route 2. As expected the TLC approach starts forwarding all packets on route 2 within one packet transmission time. The FBC case responds more slowly, as only new flows are forwarded on the route 2 while existing label switched flows will still follow route 1. The results for FBC are significantly worse than we expected, as discussed in Section 4. In the best case (June 1995 trace) 50% of packets are forwarded on the new route (route 2) 150 seconds after the route change while 85% is reached after 1350 seconds. The November 1997 trace

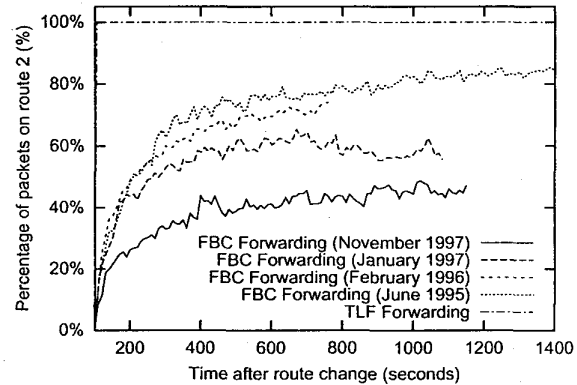


Fig. 3. Time response to a single route change

did not even reach 50% of packets on route 2 when the trace ended 1200 seconds after the route change. In order to explain these results it is necessary to examine the trace characteristics.

Figure 4 shows the characteristics of the traces in more detail. Figure 4 (a) to (d) compares the percentage of packets on route 2 after the route change with the percentage of total flows for each trace. A significant difference can be seen between the percentage of flows and the percentage of packets in each case, particularly the later traces. Flow length statistics are shown in Figure 4 (e) to (h). These graphs show the contribution of different length flows to the total number of cells and flows. We can see that each trace has a large percentage of short flows which contribute a significant proportion of packets, and a very small proportion of large flows that also contribute significantly to the total percentage of packets. This characteristic gets more exaggerated in the later traces particularly the 20/11/97 trace. In this trace only 5% of flows exceed 1150 seconds, however, this small percentage of flows contributes over 50% of total cells. In the best case the 21/06/95 trace 2% of flows are longer than 1550 seconds and contribute 18% of cells.

5.1 Altering Flow Detection

In the previous section we show that the poor sensitivity of FBC forwarding is linked to a small number of long flows that contribute a significant proportion of total cells. In this section we examine possible ways of improving the sensitivity of FBC forwarding to route table changes. We examine two possible alternatives: reducing the flow timeout value to below 60 seconds; and fixing an upper bound on the flow length.

A flow timeout of 60 seconds is commonly used for FBC forwarding [2]. If the flow timeout is reduced then the number of long flows will also be reduced improving sensitivity to route changes. However, this will also have

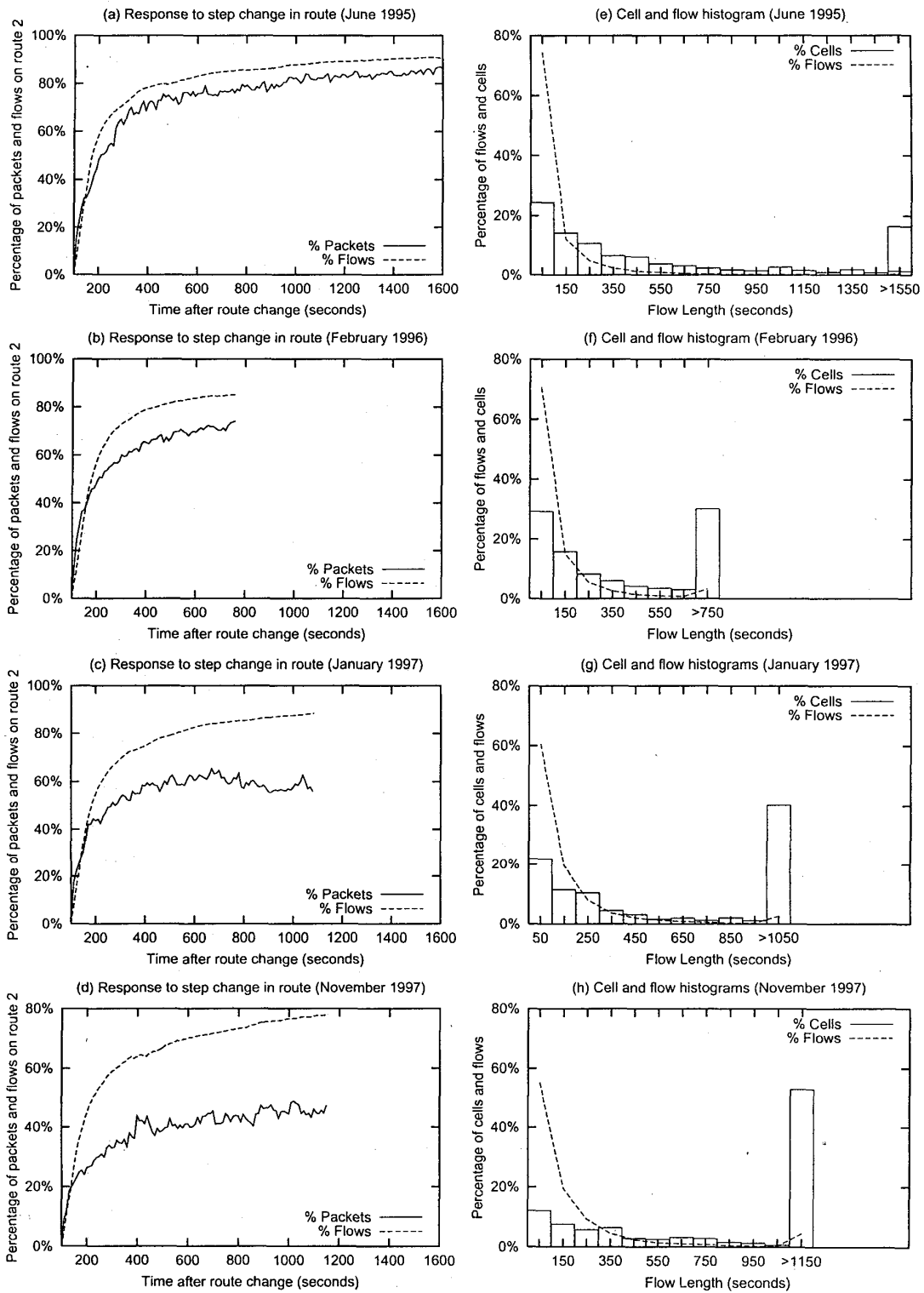


Fig. 4. Flow characteristics

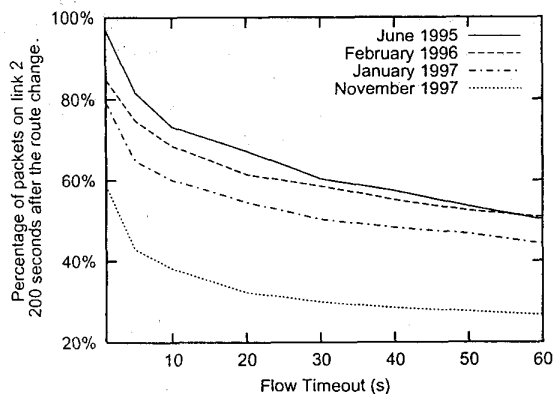


Fig. 5. Effect of varying flow timeout on sensitivity to route change

the undesirable side-effect of increasing the percentage of packets forwarded at the network layer. We examine the effect of reducing the flow timeout value by repeating the experiment in the previous section with flow timeouts of between 1 and 60 seconds. The results of this experiment can be seen in Figure 5 which shows the percentage of packets on link 2 (200 seconds after the route change) versus flow timeout. In the worst case (trace: 20/11/97) the effect of reducing the flow timeout from 100 seconds to 10 seconds is an increase in percentage of packets from 20% to 40% on the new route 200 seconds after the route change. The other traces experience a similar magnitude improvement. The penalty of this performance improvement is an increase in the percentage of packets that require network layer forwarding. This increase can be seen in Figure 6 and becomes quite significant at flow timeouts below 10 seconds. In the worst case example (the January 1997 trace) the network layer forwarding requirement for a flow timeout of 60 seconds was 10.6% this increases by 7% to 17.5% when the timeout is reduced to 10 seconds.

We now investigate our proposal of limiting the maximum length of flows to improve the sensitivity of FBC forwarding to route table changes. Using the flow length characteristics shown in Figure 4 we predict that implementing a Maximum Flow Length (MFL) will substantially improve sensitivity while resulting in a minimal reduction in the percentage of packet switched. When this flow length is reached the cut-through created for the flow is ended, and re-created, if necessary, according to the normal flow detection criteria. We examine the performance of FBC forwarding with varying MFL. Figure 7 and Figure 8 show the real-time response to a single route change with an MFL of 100 seconds and 200 seconds respectively. With an MFL of 100 seconds 50% of packets are forwarded on route 2 after 10-30 seconds. With an MFL of 200 seconds 50% of packets are forwarded on route 2 af-

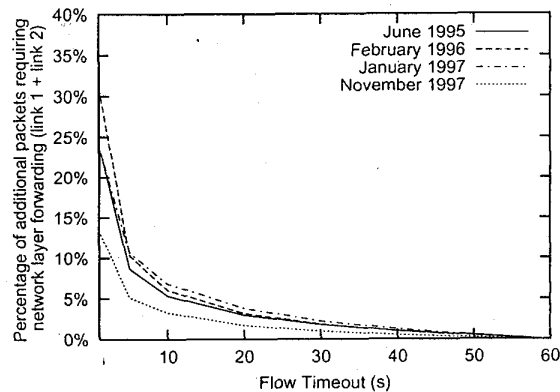


Fig. 6. Effect of varying flow timeout on the percentage of packets requiring network layer forwarding

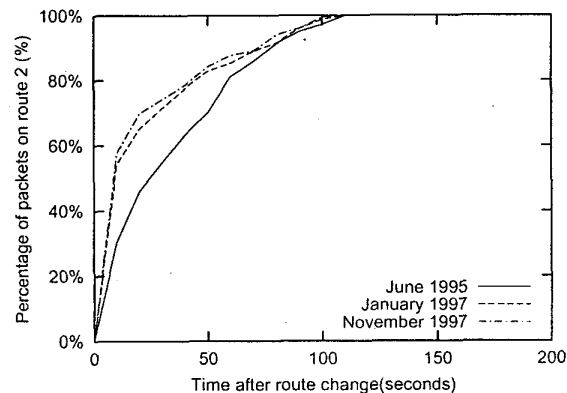


Fig. 7. Time response to a route change for MFL=100

ter 10-105 seconds. The cost of this performance increase is a reduction in the percentage of packets forwarded on cut-through routes leading to a corresponding increase in the number of packets forwarded at the network layer. The reduction of packets forwarded on cut-through routes introduced by use of an MFL can be seen in Figure 9. With an MFL of 100 the cost is a 2%-3.4% reduction of packets on cut-through routes. An MFL of 200 leads to a significantly lower penalty of between 1% and 1.5%. In the worst case example (January 1997 trace) the introduction of an MFL of 100 increases the network layer forwarding requirement by 1.5% (from 10.5% to 12%) while reducing the response time to a route change from 400 seconds to approximately 10 seconds.

6. Conclusions

In this paper we described a series of experiments designed to investigate the performance differences of the use of table-linked cut-through and flow-based forwarding

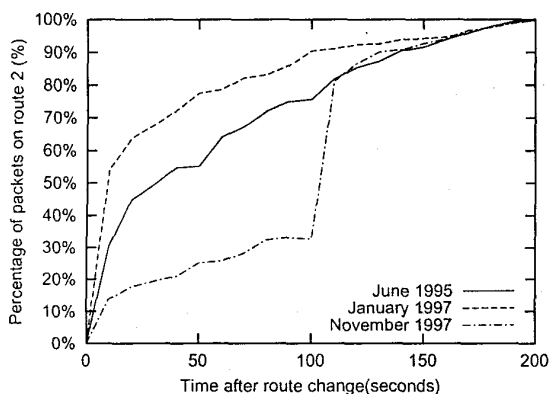


Fig. 8. Time response to route change for MFL=200

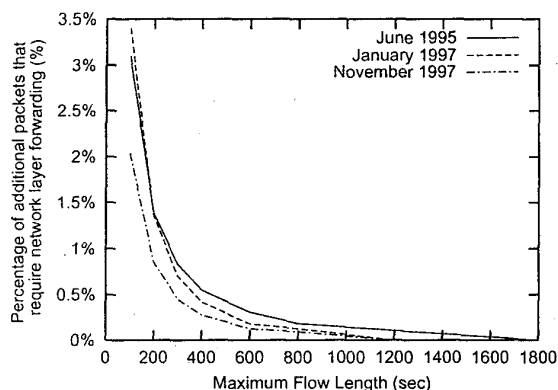


Fig. 9. Network layer forwarding penalty versus maximum flow length

in an environment where routing tables change frequently in response to changing network congestion. Firstly, this paper examined time response curves depicting the reaction to a single change in the lowest-cost route using several traffic traces. Flow-based cut-through forwarding was shown to react significantly slower than expected, by examination of published flow length distributions [6], to changes in routing tables. The time lapse varies between 140 and over 1200 seconds (depending on trace used) before more than 50% of flow-based packets are forwarded on the new route after a route change. This problem was seen to get progressively worse between June 1995 and November 1997 indicating a possible trend in traffic characteristics. The reason for this unexpectedly slow reaction to route changes was shown to be caused by a small percentage (between 3% and 4%) of long flows that contribute significantly to the total packet rate (18-50%).

Flow detection parameters for flow based cut-through forwarding were examined to determine their effect on the sensitivity to route changes. We found that reducing the flow timeout parameter improved performance slightly if a flow timeout of 10 seconds or less was selected. However, this also results in a significant reduction in the percentage of packets forwarded over label switched paths and therefore an increase in the network layer forwarding required. The introduction of a maximum flow length was shown to result in a significant improvement in the sensitivity to route changes. Selecting a maximum flow length of 200 seconds reduces the number of packets on cell level cut-through paths by only 1.5% in the worst case examined. This work currently being extended to examine the effect of aggregation.

References

- [1] R. Callon, P. Doolan, N. Fieldman, A. Fredette, G. Swallow, and A. Viswanathan. A framework for multiprotocol label switching. Internet draft, IETF, November 21 1997. Expires May 21 1998.
- [2] Peter Newman, Greg Minshall, and Thomas Lyon. Ip switching - atm under ip. *IEEE/ACM Transactions on Networking*, 6(2):117-129, 1998.
- [3] Hassan M. Ahmed, Ross Callon, Andrew Malis, and John Moy. Ip switching for scalable ip services. *Proceedings of the IEEE*, 85(12):1984-97, 1997.
- [4] Yakov Rekhter, Bruce Davie, Eric Rosen, George Swallow, Dino Farinacci, and Dave Katz. Tag switching architecture overview. *Proceedings of the IEEE*, 85(12):1973-1983, 1997.
- [5] Curtis Villamizar. Ospf optimized multipath (ospf-omp). Internet Draft draft-ietf-ospf-omp-01, Internet Engineering Taskforce, October 9 1998.
- [6] Kimberly C. Claffy, Hans-Werner Braun, and George C. Polyzos. A parameterizable methodology for internet traffic flow profiling. *IEEE JSAC*, 13(8):1481-1493, 1995.
- [7] Andre Fredette. *Multi-Protocol Over ATM Version 1.0*. The ATM Forum, 1997.
- [8] P Doolan, B Davie, D Katz, Y Reckhter, and E Rosen. Tag distribution protocol. Internet Draft draft-doolan-tdp-spec-01.txt, IETF Network Working Group, May 1997.
- [9] David Glazer and Carl Tropper. A new metric for dynamic routing algorithms. *IEEE Transactions on Communications*, 38(3):360-367, 1990.
- [10] R Coltun. The ospf opaque lsa option. Request For Comment RFC2328, IETF Network Working Group, July 1998.