Packets or ports - the decisions of IP switching

Mika Ilvesmäki [a], Kalevi Kilkki [b] and Marko Luoma [a]

[a] Helsinki University of Technology, PL 3000, 02015 TKK, Finland
[b] Nokia Research Center, Helsinki, Finland

ABSTRACT

One of the most important network elements in the Internet are the routers which do relaying of IP packets. Because of growth of the Internet routers currently experience serious problems in relaying traffic in a satisfying speed. The idea of switching Internet traffic flows has recently been introduced and a new technology called IP switching has emerged. Several differing technological solutions have been suggested. In this paper we describe and compare two methods in flow-based IP switching to make the decisions whether to switch internet traffic flows to separate ATM-connections. Traffic measurements are made in two networks of varying size and based on a specific three-stage flow analysis we suggest that the decision to switch should be made as flexible as possible due to the expected diversity of traffic profiles in different parts of the network. This way the optimal service cluster could be switched and router resources could be optimally utilized. A simple model to determine workload to an IP switch is introduced. Using this model we see that the workload of the flow setup component and the routing component may be optimized, if we use flexible methods to determine the flows that are to be separately switched.

Keywords: IP switching, Internet, routing, traffic measurements, flow switching

1. INTRODUCTION TO IP SWITCHING

The ongoing growth of the Internet requires better performance from the routers doing relaying of the internet protocol (IP) datagrams. Routers are an essential part in internet and their limited ability to perform is creating bottlenecks in the Internet. Current routers make forwarding decisions for each packet separately, but it would obviously be more efficient to make this decision only once to a series of packets, or a flow, traveling to the same destination. This is especially true if a flow of packets sent to the same destination from one source, contains a large number of packets that are being sent during a relatively long time. A number of solutions have recently emerged to deal with this issue. These solutions include Ipsilon's IP switching, Cisco's Tag switching, Toshiba's Cell Switch Router (CSR), Telecom Finland's Switching IP through ATM (SITA) and IBM's Aggregate Route-Based IP switch (ARIS). Also the Internet Engineering Task Force's (IETF) Multiprotocol Label Switching (MPLS) workgroup is studying the subject. The emerged technological solutions, although different from each other, all aim to offer in one way or the other the flexibility of routing combined with the speed of asynchronous transfer mode (ATM) switching. In this paper we concentrate on some specific issues concerning flow based IP switching. A simplified illustration of the idea of an IP switch is presented in Figure 1.

Further author information -
M.I: (correspondence): Email: lynx@luuri.hut.fi; WWW: http://keskus.hut.fi/tutkimus/ipana/mika.shtml; Fax: +358 9 451 2474
K.K: Email: kalevi.kilkki@research.nokia.com
M.L: Email: mluoma@luuri.hut.fi; WWW: http://keskus.hut.fi/~mluoma; Fax: + 358 9 451 2474

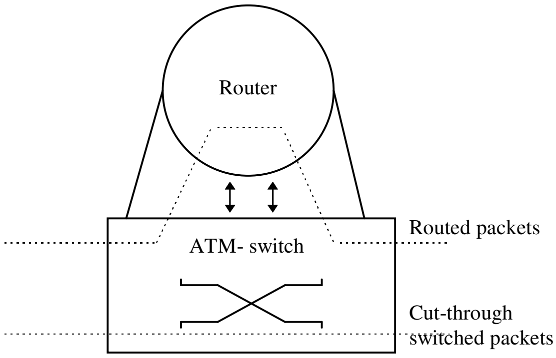


Figure 1: IP switch

The main idea behind IP switching, besides reducing the workload of an IP router, is to provide more bandwidth and a dedicated quality of service (QoS) to applications that need it the most. In the current Internet world these applications may be defined by transmission control protocol (TCP) or user datagram protocol (UDP) port numbers.
IP switching is based on detecting IP flows. An IP flow is a series of IP packets that share some common properties, usually the IP address and perhaps also the TCP/UDP-port number. This concept of an IP flow is illustrated in figure 2.

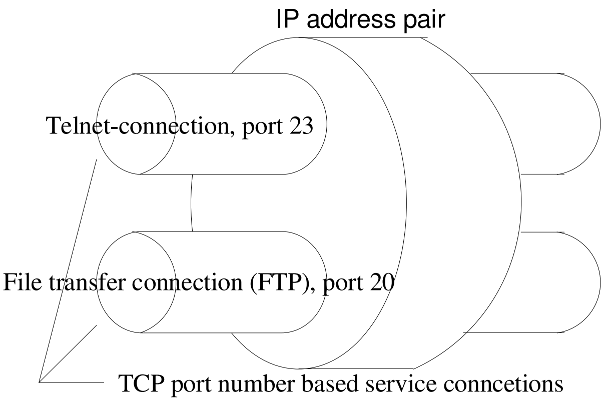


Figure 2: IP flows

A new flow is established every time a packet containing a unique IP address pair and possibly a port number pair is detected. In this paper the packets belong to the same address based flow if a similar IP address pair is detected and if the time between two packets having the same address pair is less than 60 seconds [1]. An extension to this method, also discussed in this paper, is to check, besides the address pair and the 60 second time limit, the similarity of TCP/UDP-port numbers in consecutive packets to determine whether a new flow should be established or an old one be used. This latter method is also illustrated in figure 2 where different TCP/UDP-port numbers to an equal IP address constitute different flows. Different IP flows might be considered to require different QoS, e.g. error ratio, transfer delay etc. However, QoS has not been realized in current networks and practically every service used in the Internet receives now only best-effort QoS. However, it seems that more and more Internet applications demanding some degree of QoS will emerge and therefore some solution to guarantee QoS in the Internet is required.
In ATM virtual paths and channels are used to manage the available bandwidth and provide different connections different kind of QoS. This concept is illustrated in figure 3.

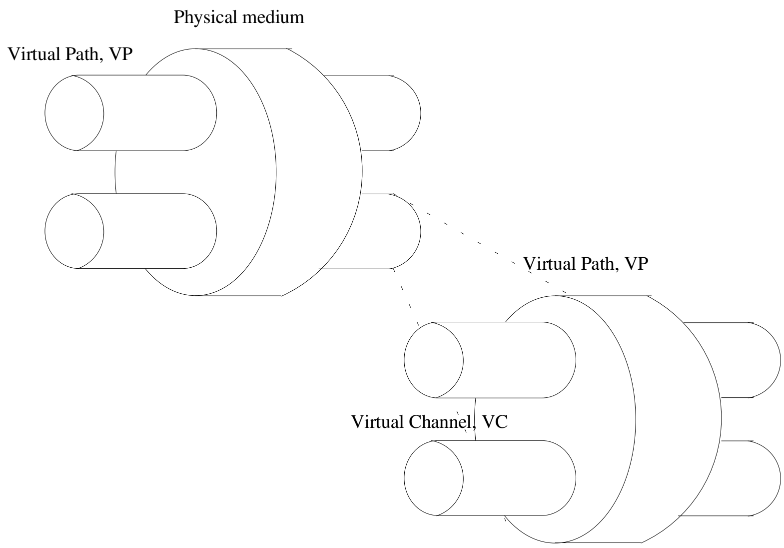


Fig 3. Virtual paths (VP) and virtual channels (VC) in ATM networks

IP switching aims to decrease the workload of routers by assigning long-lived IP flows which contain a relatively large amount of data to transfer to separate virtual channels. This way it is possible to assign certain services to separate VP/VC's and offer them at least a better QoS than can be realized in a default connection. In this paper we examine two actual networks of different size and the effect to the estimated workload of a combined router and switch, IP switch, when IP switching is applied. Suggestions are made on what quantities should the decisions whether to switch a certain flow or not be based on.

2. A SIMPLE METHOD TO ANALYZE THE WORKLOAD OF AN ROUTER AND AN IP SWITCH

In the following a simple method to estimate and analyze the workload of an IP router relaying and establishing IP flows is presented. After having obtained a traffic trace we determine the basis for establishing a flow. These include the address pair (AP) method or the port pair (PP) method and the option to use packet count thresholds in either of these. The AP and PP method are further clarified in chapter 4 in this article. After having determined the number of packets to be switched and routed and the number of flows established, we may determine the workloads of different components in the IP switch.

To determine the workload of the combined router and flow classifier we denote the number of packets to be routed as n_r and the number of packets required to make switching of the flow possible (establishing a flow) as n_{est} and the number of flows determined from the traffic trace as n_{flow}. The total work done, W_{router}, is the amount of packets routed, and formed and sent to establish flows. Thus, the total workload of a router routing n_r packets and classifying n_{flow} flows at the same time is given in (1).

$$W_{router} = \frac{n_r + n_{flow} n_{est}}{T} , \tag{1}$$

n_r = number of routed packets
n_{est} = number of packets needed to establish a single flow
n_{flow} = number of flows to be established in the traffic trace
T = the duration of the trace

If we apply IP switching and we denote the amount of packets switched as n_{sw}, we may determine the total workload of an IP switch $W_{IPswitch}$ as (2).

$$W_{IPswitch} = W_{router} + \frac{n_{sw}}{T} , \tag{2}$$

n_{sw} = number of packets switched

In (2) we look at both the workloads of routing function and flow establishment function. This is to take into account the possible negative effects of trying to take too much work out of the routing function. This way we would only relocate the problem, not solve it.

In this work we concentrate on minimizing the $W_{IPswitch}$ as a function of n_r (number of packets routed) and the packet count threshold. We also somewhat vary the values of n_{est}, or the number of packets it takes to establish a flow.

The model presented here is an extremely simple one, not intended to provide one exact result of router and flow setup workloads but to provide one with basic knowledge about the direction of the consequences when applying IP switching in a router. If one wishes to develop the model further on, the first step should be to take into account the differing workloads that routing, flow classifying and switching result in the IP switching system.

3. TRAFFIC MEASUREMENTS

Traffic measurements were made in Helsinki University of Technology on two bridged 10 Mbit/s Ethernet local area networks during (9 a.m. - 11 a.m.) office hours. The smaller network is the local network of the laboratory of Telecommunications technology having some tens of users and the bigger one serves as the backbone network of the department of electrical and telecommunications engineering serving some hundreds of users. Both measurements resulted information of TCP/IP- or UDP/IP- packets. In addition to a timestamp, the source and destination addresses and ports were obtained from the traffic trace. The traces were obtained by using TCPDUMP. As the analysis was performed all sensitive information was removed from the traces.

4. ANALYSIS

The analysis of the network measurements is performed in three stages. In the first stage we perform a simplified flow analysis where detecting a new IP address pair in an IP packet results in creating a new flow. Also the time between two packets in a flow may not exceed 60 seconds or otherwise the flow is deleted. We then observe the proportion between sent packets and detected flows as a function of the flow lifetime. Flow lifetime is defined to be the time between the first and the last packet sent in a flow. This first stage analysis provides us the basis as to whether applying flow switching is reasonable in the network measured.

In the second stage of the analysis we observe the statistics of services relayed in the network. This is done by applying the methodology of the first stage analysis to each service (that is, to a well-known TCP/UDP-port number pair and IP address pair) detected in the trace. This second stage analysis points out which services form the majority of packets and flows and would therefore possibly be suitable for switching.

The third and the final stage is to determine the effect to the routing and flow setup component of an Internet router when using packet count thresholds before the router is to establish either address pair (AP) based or address and port pair (PP) based flows. This is done by using the information gathered in the stage one and two analyses and applying the model for determining the workload of an IP switch introduced earlier in this paper. The method of analysis presented here has some similarities to a traffic analysis presented in [3] so results obtained here may be compared.

4.1 Flow analysis

The measurement of the smaller network contained approximately 143 000 packets and the measurement lasted approximately 5038 seconds (approximately 1 h 45 minutes). In this time a total number of 1402 flows would be created resulting in an average speed of 0.28 flow setups/s. At the first stage only IP address pair based flows were studied. The result in the smaller network is illustrated in figure 4.

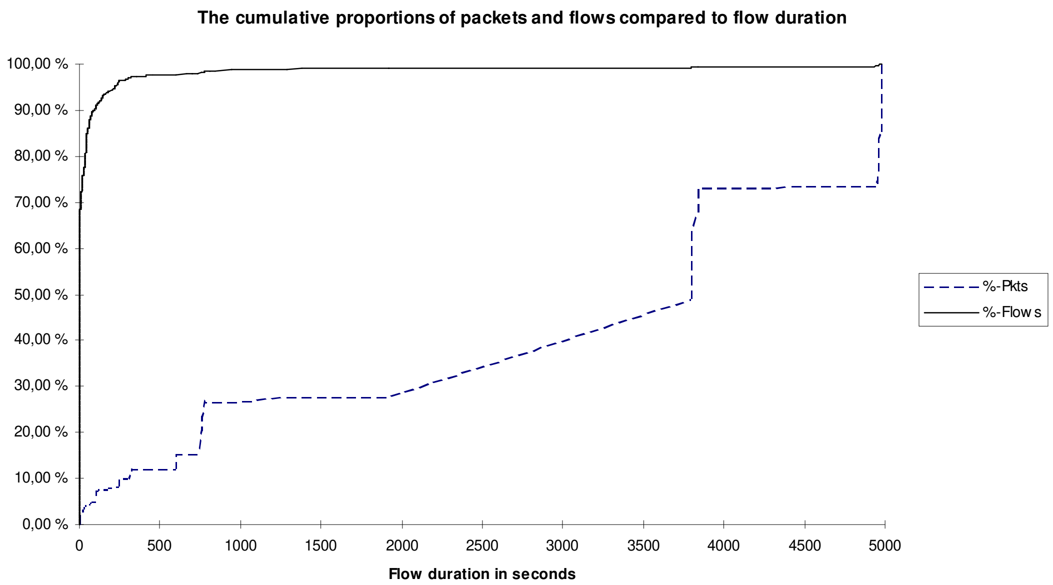


Figure 4: The flow analysis for the small network

Figure 4 shows that in the smaller network roughly 50 % of flows contain only a little over 1 % of packets transmitted but that 10 % of the longest flows contain approximately 95 % packets.

The measurement of the larger network contained a little over 1 100 000 packets in approximately 3600 seconds (1 hour). The number of flows established was 7238 resulting in an average of 1.97 flow setups/s. Here too at the first stage only flows and amounts of packets based on IP address pairs were studied. The result is illustrated in figure 5.

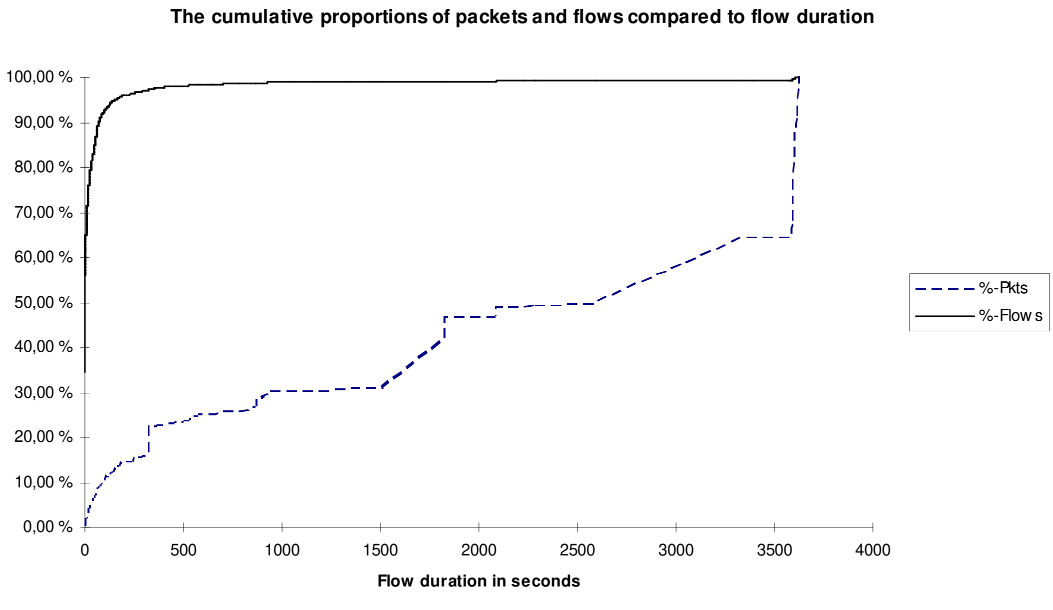


Figure 5: The flow analysis of the larger network

Figure 5 shows that in the larger network roughly 70 % of flows contain less than 3 % of packets transmitted but that 10 % of the longest flows contain approximately 91 % packets. These results are according to the measurement made in the smaller network.

From the first stage analysis we can conclude that flows with a relatively long duration (that is, relative to the measurement interval) seem to carry the majority of packets. This suggests that long-duration flows may be switched and that if done so this would substantially reduce the workload of the routing component of an Internet router.

4.2 Protocol analysis

To achieve clarity and simplicity in viewing the results from the second stage analysis only those well-known TCP/UDP-port numbers and the port number space reserved for X11-connections are presented in tables 1 and 2 and are considered suitable for switching . The port number has been included in the table if its total share of packets in the trace is more than approximately 0.05 %. This is an arbitrary limit and may be freely adjusted.

When the second stage protocol analysis, where flows are created based on both IP address pairs and TCP/UDP-port numbers, is applied to the measurement results from the smaller network we see that almost 58 % of the packets are suitable for switching (Table 1) and these packets form 56 % of all flows detected. Therefore we would be able to switch nearly half of the packets with reasonable load from switching the flows. The actual rate of flow setups when switching only those TCP/UDP-protocols included in Table 1 is 1.6 flow setups/s compared to the maximum of 2.8 flow setups/s if every flow detected in this second stage analysis would be switched. So using protocol analysis and picking out the ones with the most favorable qualities regarding switching (i.e. long duration flows with many packets) gives us about 40 % lower rate of switching compared to the process where every detected flow would be switched.

When the second stage protocol analysis is applied to the measurement results from the larger network we see that almost 51% of the packets are suitable for switching (Table 2) and these packets form a little under 47 % of all the flows detected. Therefore we would be able to switch half of the packets with reasonable load to the flow setup component. The actual rate of flows switched is 13 flow setups/s compared to 28 flow setups/s if every flow detected in this second stage analysis would be switched. So using protocol analysis and picking out the ones with most favorable qualities regarding switching gives us nearly 55 % lower rate of switching compared to the process where every detected flow would be switched.

The results from second stage analysis result in somewhat higher rates of flow setups than in the first stage analysis but using TCP/UDP-port pairs we are able to achieve more granularity in switching and thus enabling to offer QoS to those connections really needing it. Also using the TCP/UDP-port pairs we are still able to significantly reduce the workload of the router component. Therefore, we should be picking up suitable protocols from tables 1 and 2 and then be switching them. However, when comparing tables 1 and 2 they indicate that even two closely related networks result in somewhat different service profiles (that is port-number profiles). Furthermore, if we compare the service-profiles observed in other similar studies[3,4] we see that to pick up a stable set of protocols to be switched may easily result in an non-optimal service-cluster to be switched in different network environments.

An alternative method, not dependent on changes in service profiles, should be considered. In addition it should be noted from tables 1 and 2 that the currently widely used http-protocol does not necessarily qualify in the top protocols to be switched because of its feature to reserve new port numbers (and therefore to establish new flows) for different elements in an HTML-document. This detail is to be fixed in the next version of the http-protocol and will result the protocol to be a bit more suitable for switching. Furthermore, it should be noted that the demand in secure network connections results distorting the service profile. For instance, the use of SSH (secure shell, port 22) hides very effectively the actual services used. This is an issue that needs further investigation.

Table 1: The results of flow analysis on the smaller network

Protocol	Port	% flows	% pkts	flows/s	pkts/s
x11	6000	4,06 %	37,54 %	0,11	22,68
netbios-ns	137	4,09 %	8,38 %	0,11	5,07
ssh	22	1,85 %	3,03 %	0,05	1,83
www-http	80	5,04 %	2,07 %	0,14	1,25
reserved	1023	0,17 %	1,79 %	0,00	1,08
domain	53	29,37 %	1,36 %	0,82	0,82
netbios-dgm	138	7,26 %	1,01 %	0,20	0,61
unassigned	1013	0,02 %	0,45 %	0,00	0,27
unassigned	1005	0,01 %	0,45 %	0,00	0,27
unassigned	1017	0,05 %	0,39 %	0,00	0,23
bootpc	68	0,01 %	0,26 %	0,00	0,16
bootps	67	0,01 %	0,26 %	0,00	0,16
reserved	0	0,93 %	0,18 %	0,03	0,11
device	801	0,93 %	0,18 %	0,03	0,11
auth	113	0,04 %	0,12 %	0,00	0,07
efs/router	520	0,03 %	0,12 %	0,00	0,07
unassigned	1021	0,21 %	0,08 %	0,01	0,05
unassigned	1014	0,02 %	0,07 %	0,00	0,04
pop3	110	0,87 %	0,07 %	0,02	0,04
unassigned	1022	0,13 %	0,07 %	0,00	0,04
hosts2-ns	81	0,20 %	0,04 %	0,01	0,03
sunrpc	111	0,78 %	0,04 %	0,02	0,03
Total		**56,08 %**	**57,96 %**	**1,56**	**35,02**

Table 2: The results of flow analysis on the larger network

Protocol	Port	% flows	% pkts	flows/s	pkts/s
www-http	80	8,26 %	30,85 %	2,28	171,12
x11	6000	0,26 %	4,57 %	0,07	25,32
domain	53	17,38 %	4,16 %	4,80	23,08
auth	113	1,61 %	1,94 %	0,44	10,75
device	801	4,14 %	1,07 %	1,14	5,92
nntp	119	0,08 %	1,05 %	0,02	5,83
sunrpc	111	7,65 %	1,04 %	2,11	5,75
who	513	0,06 %	0,75 %	0,02	4,14
unassigned	897	0,01 %	0,72 %	0,00	4,01
ssh	22	0,44 %	0,71 %	0,12	3,92
telnet	23	0,32 %	0,60 %	0,09	3,33
ntp	123	0,97 %	0,44 %	0,27	2,42
https	443	0,27 %	0,38 %	0,07	2,12
unassigned	1021	0,06 %	0,37 %	0,02	2,03
smtp	25	0,22 %	0,30 %	0,06	1,66
reserved	0	0,95 %	0,23 %	0,26	1,29
unassigned	1015	0,03 %	0,20 %	0,01	1,10
ftp-data	20	0,05 %	0,19 %	0,01	1,04
finger	79	0,83 %	0,16 %	0,23	0,88
pop3	110	0,66 %	0,14 %	0,18	0,77
netbios-ns	137	1,28 %	0,12 %	0,35	0,68
ftp	21	0,06 %	0,11 %	0,02	0,62
ipx	213	0,01 %	0,09 %	0,00	0,52
reserved	1023	0,15 %	0,07 %	0,04	0,41
netbios-ssn	139	0,14 %	0,07 %	0,04	0,39
compressnet	2	0,04 %	0,06 %	0,01	0,32
bootpc	68	0,00 %	0,06 %	0,00	0,31
netbios-dgm	138	0,61 %	0,04 %	0,17	0,24
cisco-fna	130	0,02 %	0,04 %	0,00	0,24
tcpmux	1	0,05 %	0,04 %	0,01	0,20
xns-time	52	0,03 %	0,03 %	0,01	0,19
xfer	82	0,02 %	0,02 %	0,01	0,09
Total		**46,67 %**	**50,60 %**	**12,88**	**280,69**

4.3 The effect of packet count threshold to flows detected

Another method to base the decision to switch is counting the number of packets in a flow to be received before switching the flow. Here we aim the port pair (PP) based flow switching speed to be the same that resulted from the first stage address pair (AP) based flow switching. Using the packet count threshold in both the AP and PP methods and then using (1) and (2) we are able to see in Figure 6 that the switching threshold in PP-switching should be somewhere around 3 to 5 packets in the smaller network to achieve the same performance measured in AP-switching.

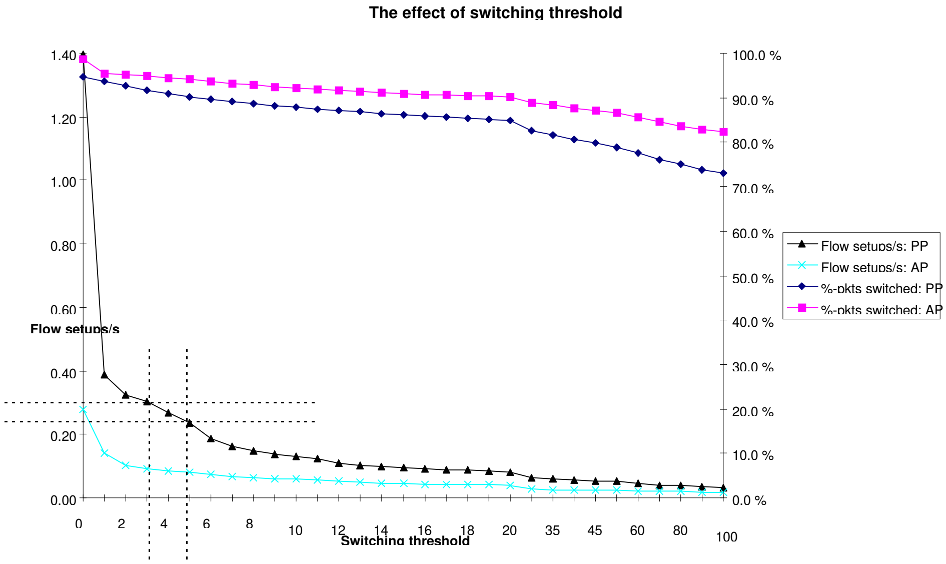


Figure 6: The effect of switching threshold in the small network (Note the nonlinear x-axis)

In the larger network we see from Figure 7 that the switching threshold in PP-switching should be roughly approximately 10-20 packets.

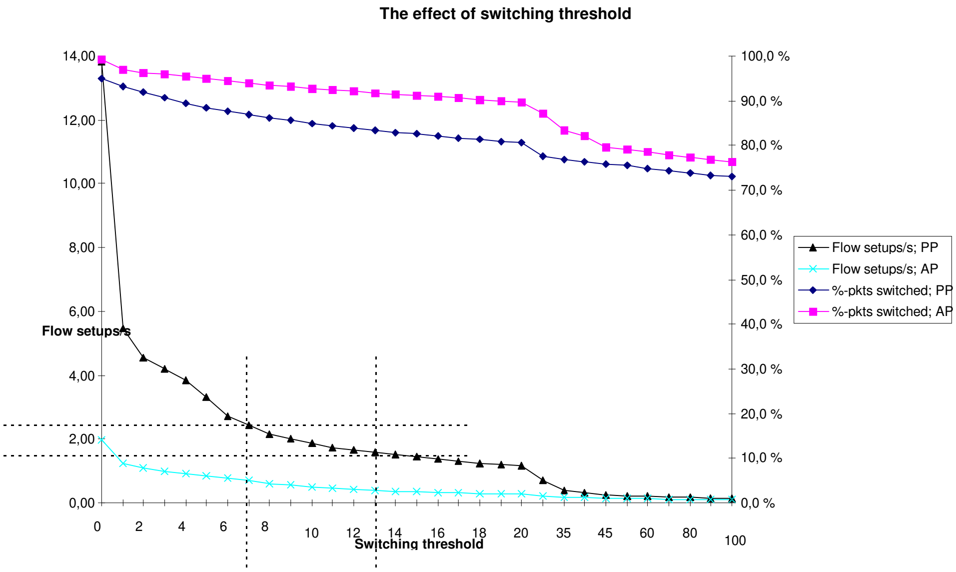


Figure 7: The effect of switching threshold in large network (Note the non-linear x-axis)

The phenomena present in figures 6 and 7 where the speed of flow setups rapidly decreases when applying packet count thresholds to PP-switching results because pure PP-switching is not able to predict that some PP-flows contain only very few packets. The introduction of packet count thresholds in PP-based flow switching partly eliminates this problem. When comparing the speeds in different methods with which flows are established, presented in table 3, we see that if packet count threshold is applied to TCP/UDP-port pair based switching a significant deduction in flow setup/s is achieved. It should also be noted that if no packet count thresholds are used then the AP-method is the one that decreases the amount of packets to be routed the most. Further on, we still need to assess the workload of the routing component and setting up flows.

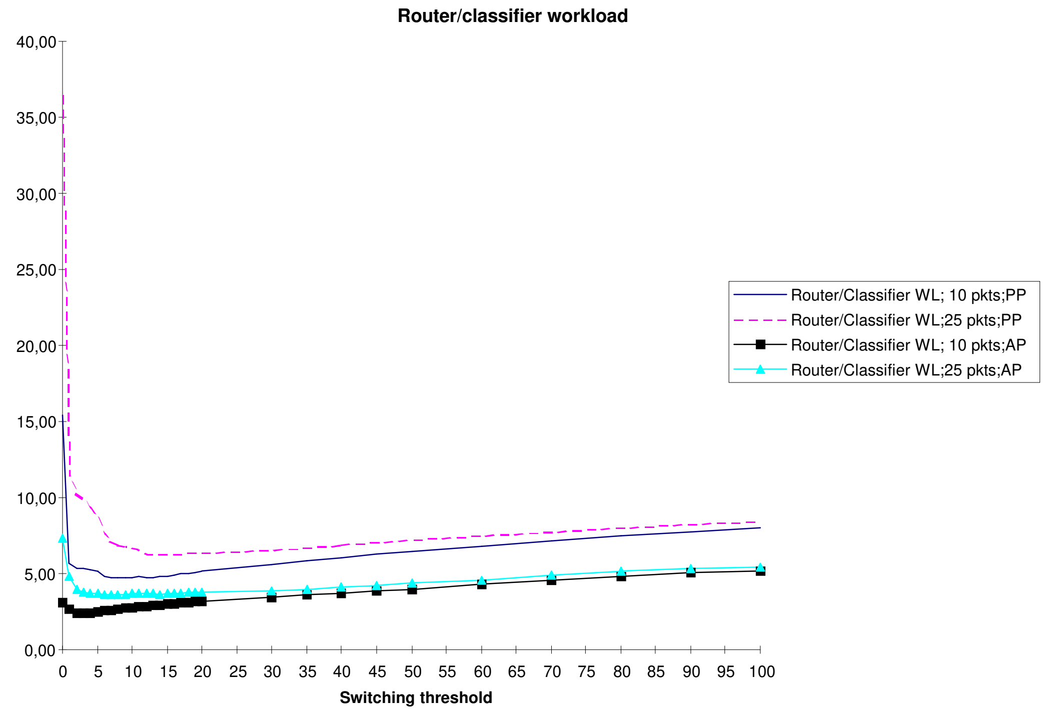
Table 3: Comparison of flow setups/s

Network type	Address Pairs	Port Pairs	PktThreshold of 10 pkts
Large network	1.97	28	≈ 2.0
Small network	0.28	2.8	≈ 0.1

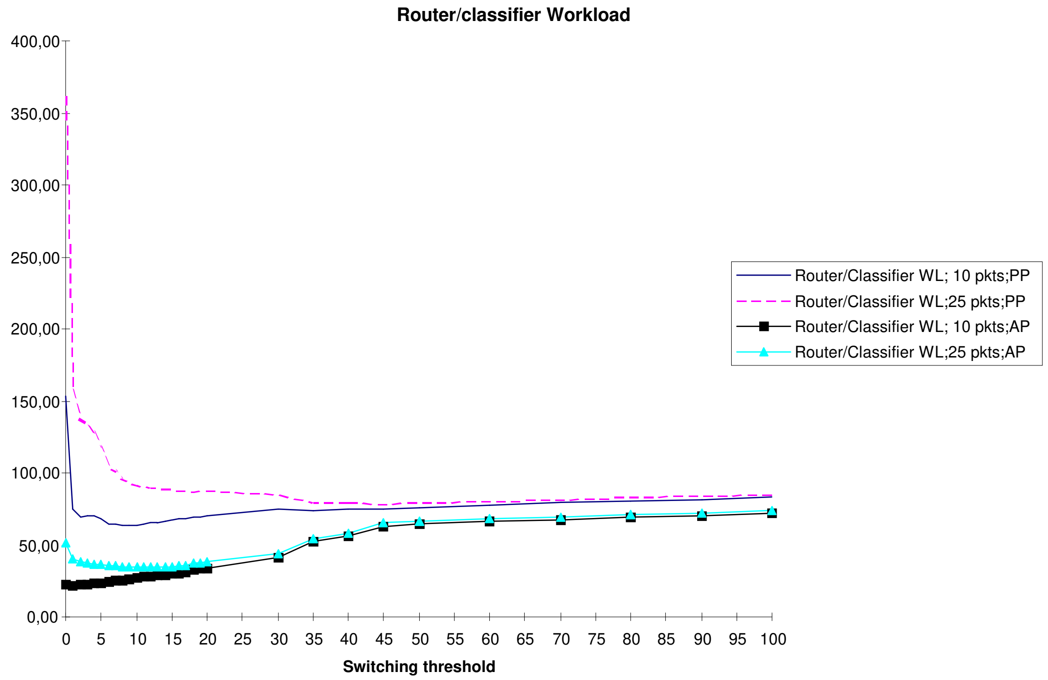
4.4 Routing and classifying flows

Using (1) and varying the value of n_{est} (the packets needed when establishing flow) between 10 and 25 packets and applying the data gathered in previous measurements we get the workload of the routing and flow classifying component in both networks illustrated in figures 8a) and 8b). In both of the figures 8a) and 8b) the effect on the workload is illustrated as a function of switching threshold. We see that regardless of how many packets are required to switch a flow approximately 10 to 15 packets is the amount to be used as the packet count threshold if a notable decrease in the router and flow classifier workload is desired. Especially it should be noted that the achieved decrease in the router's/classifier's workload depending on the size of the network is over 60 % and can be estimated in this experiment to rise to as high as 70%. This is achieved when TCP/UDP-port pair based flow switching combined with packet count thresholds per flows is used. The results indicate that the application of switching the IP flows would offer significant decrease in router/classifier workload.

When using the AP-method to establish flows we achieve the greatest decrease in router/classifier workload. However, when using the AP-method we can not determine QoS to separate services used. Still, it might preferable to be able to offer a dedicated QoS to specified services only and it is our recommendation that the PP-method together with packet count thresholds is the way that the flows to be switched should be determined. However, an IP switch has also the flow setup component and its performance should be noted too.



a) smaller network

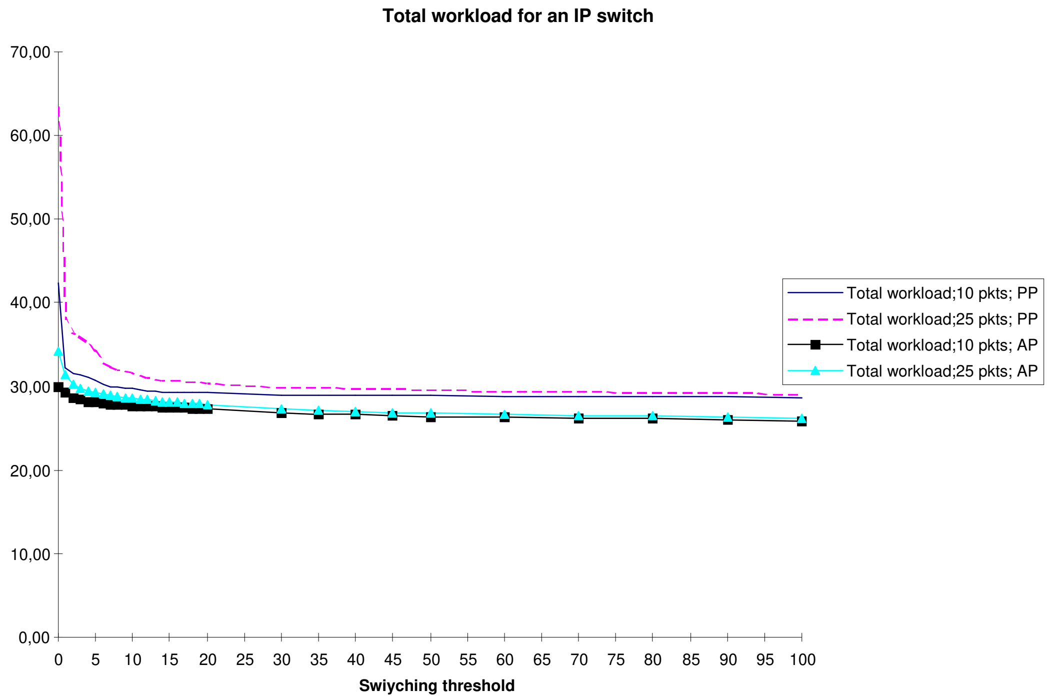


b) bigger network

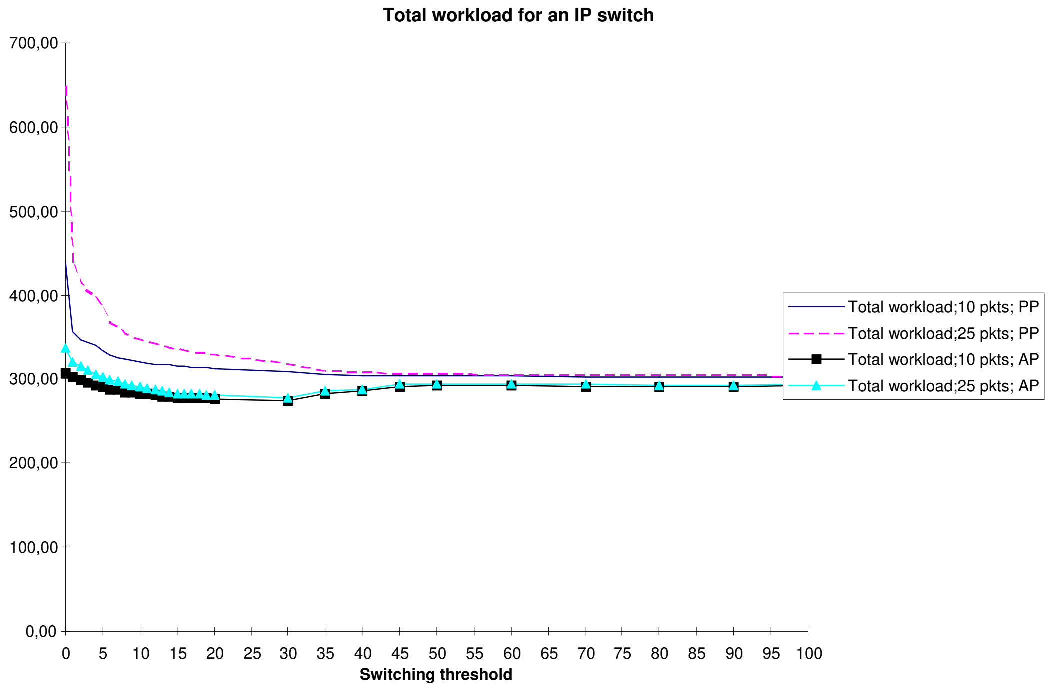
Figure 8: The combined workload of the router and flow classifier in the a) small and b) bigger network

4.5 Routing and switching combined

If we combine the workloads of the routing, flow classifying and switching using (2) to calculate the total workload of an IP switching system using packet count thresholds. The results are illustrated in figures 9a) and 9b).



a) smaller network



b) bigger network

Figure 9: The combined workload of routing and switching in a) small and b) bigger network

If we take into account the combined effect of routing and switching the flows we see that the optimal packet count threshold is somewhere between 10 to 25 packets depending on how much we want to optimize the total workload of the IP switch.

5. CONCLUSIONS

In this paper we have assessed the applicability of the IP switching concept and the amount of work done in different components of an IP switch. The evaluation is based on traffic measurements in two university networks: the smaller network is a typical local area network with some tens of users, and the larger one is a backbone network with some hundreds of users. The main goal of this study has been to develop simple and efficient algorithms for the decision when an IP flow should be switched to a separate ATM-connection. Emphasis has also been in the optimal use of resources of Internet routers and, on the other end, the ability to offer QoS to the Internet services used.

Even the results of the flow analysis of two networks, although closely connected with each other, give cause for slightly different conclusions regarding the service profile present in these networks respectively . This indicates that it will be very difficult to develop any general and accurate rule for the decision if using the TCP/UDP-port numbers as the basis for the decisions. In this light it is of great importance to be sure that the decisions to switch are not only based on assumptions of the service profiles, but a more reliable and general method should be used. Taking into account the different and dynamic traffic profiles used in the internet the decisions to switch an IP flow should, in our opinion, be based on both the TCP/UDP-port numbers and the number of packets received on the flow. This allows for more adaptability and dynamic use of resources in cases when the workload of an IP switch increases. With this kind of scheme it is possible to decrease the workload of the router at least up to 70 %. Using packet count thresholds and port numbers as the basis for flow setup decisions has the advantage of giving QoS to the connections actually needing it, i.e. to those flows which contain a substantial number of packets. However, if the sole purpose of IP switching is to relieve the workload of routers then the only answer is to base the switching decisions on IP address pairs.

If a static choice of TCP/UDP-ports is used as basis for flow switching, we restrict ourselves to a particular traffic profile that evidently, also according to this and some other studies, changes in time or space. Also the use of pre-determined TCP/UDP - ports as basis for switching decision blinds the router from flows that are usually long lived but happen to be short lived at times. One alternative might be to use real-time or near-real-time measurements of network traffic to establish dynamically the applications and TCP-port numbers, i.e. the service cluster, which should be switched. This could, however, lead to inconsistency of switched protocols in large networks.

The subject of real-time measurements in IP switching environments and the determination of the service profile, or service cluster, that is considered possible to be switched are important subjects in IP switched networks and should be carefully studied in the future. We further suggest that particularly the methods and applications of neural networks, intelligent agents, pattern recognition and digital signal processing should be studied.

The general decrease of workload and increase on the performance of an IP switch is furthermore dependent on several different factors mentioned, but not studied to the depth in this work. These factors include the time limit before creating a new flow and the parameters that able the flow to be switched. The analysis presented in this work should be carried further on by varying the aforementioned parameters. Furthermore, the relation between creating ATM-connections and routing packets and their effect on the total performance of the system needs further studies and the workload model presented here should be further developed to be more accurate and precise, and to more resemble the real IP switch implementations. The present workload model only gives indications of those phenomena when applying IP switching.

ACKNOWLEDGMENTS

This work was carried out in the laboratory of Telecommunications Technology in Helsinki University of Technology as a part of the IPANA-project (IP Atm Network Architectures).

REFERENCES

1. K. Claffy, H-W. Braun and G. Polyzos, “A parameterizable Methodology for Internet Traffic Profiling”, IEEE Journal on Selected Areas in Communications, Vol. 13 No. 8 pp. 1481-1494, 1995.
2. P. Newman, G. Minshall, T. Lyon, L. Huston, “IP switching and Gigabit Routers”, IEEE Communications Magazine, Vol. 35 No. 1, pp. 64-68, 1997.
3. G. Minshall, T. Lyon and P. Newman, “Flow Labeled IP: Connectionless ATM under IP”, Ipsilon Networks Inc., USA. 1996.
4. S. Lin and N. McKeown, “A Simulation Study of IP switching”, Technical Report CSL-TR-97-720, Stanford University, USA, 1997.