

Flow Switching: To Switch or Not to Switch *

Peter Newman, Tom Lyon, and Greg Minshall[†]

Ipsilon Networks Inc.[‡]

Growth of interest in the use of ATM for switching IP traffic has once again raised the specter of connectionless versus connection-oriented transfer. In this brief note we report some results in which we investigate the benefit of classifying IP flows into those that have traffic characteristics for which it is worth the effort of establishing an ATM connection and those that are better handled by connectionless forwarding.

We obtained a traffic trace from the Internet backbone. The trace contains five minutes of traffic taken on Sep 25, 1995. It was taken by monitoring an FDDI ring that connects traffic from the San Francisco Bay Area to and from the Internet backbone. The trace includes a timestamp, IP source and destination addresses, the packet length, and source and destination port numbers for each packet.

We first investigated the flow characteristics of each of the protocols present in the trace. The protocol is defined as either the IP protocol number of the packet, or if that indicates TCP or UDP, the protocol number plus the well-known port number from either the source port or the destination port numbers. (If neither port number is recognized, the source port number is used.) For each packet in the trace we check to see if this is a new flow. If so, a new flow record is created, else the statistics of the existing flow record are updated. For this purpose two packets belong to the same flow if they have the same source and destination IP addresses and the same protocol. A flow is deleted after it has remained idle for 60 seconds although the duration of the flow is recorded as being the time from when it was created until the time of the last packet transmitted on the flow. The traffic flow analysis presented in [1] suggests that a timeout value of the order of 60 seconds is a reasonable compromise between the size of the flow table and the probability of deleting flows that will shortly become active again.

The results are presented in table 1 for all protocols with a recognizable protocol that contributed more than 0.05% of the total packets in the trace. (The table accounts for about 82% of the total number of packets.) For each protocol the table gives the percentage of the total number of flows, packets, and bytes contributed by that protocol.

It gives the mean number of new flow arrivals per second after an initial startup phase of 60 seconds, the mean number of packets/s, the average duration of each flow, the average number of packets transmitted across each flow, and the mean number of bytes per packet. It also gives the mean packet arrival rate over the lifetime of each flow averaged for all flows of the same protocol (pkts/s/flow). Protocols with characteristics for which it appears worth establishing a switched connection are marked “✓”. These are protocols with an average flow duration in excess of about 20 seconds and which transmit an average of more than about 40 packets per flow (an arbitrary decision). If we assume that these flow characteristics are a property of the application behind the protocol, and the manner in which people use the applications, then for any individual protocol the results should remain relatively independent of the position in the network that the measurement is made. The characteristics of each protocol should also change only slowly over time. Thus we can use this information in deciding whether to establish a switched connection for any given packet.

It is interesting to note that http (web traffic) shows an average of 74 packets per flow, much higher than the value typically quoted (about 15–20 packets per flow). This is because we are looking at host pair flows, which allow multiple TCP connections between the same two IP addresses to share a single flow, rather than assuming a separate flow for each TCP connection. Thus the entire conversation between a client and a particular web server is regarded as a single flow unless it times out.

In a second experiment each packet is first classified according to its protocol. If it belongs to a protocol marked “✓” in table 1 it is suitable for switching. If neither the source nor destination port numbers are well known (less than 1024 or a recognized registered number) we assume the packet is suitable for switching if belongs to TCP but not if it is UDP. (This is the best guess we can make for packets that do not have a recognizable port number.) For those packets classified for switching we check to see if a suitable switched connection exists. If the search fails a new connection is established. For convenience we ignore the time required to establish a connection. This yields an upper bound on possible performance for switching and removes the vagaries of the signalling implementation from our results. In this experiment a connection is suitable if it exists between the same source and destination IP address as the packet being processed. Connections are deleted after a timeout of 60 seconds. In this experiment 84% of the packets and 91% of the bytes in the trace are recognized as suitable for switching. A mean of 92 flows per second are established after an initial startup phase of 60 seconds, with a 95th

* Presented at the NSF Workshop on Internet Statistics, Measurement, and Analysis, San Diego, Feb. 1996.

[†]{pn,pugs,minshall}@epsilon.com

[‡]<http://www.ipsilon.com>

Protocol	port		%flows	%pkts	%bytes	flows/s	pkts/s	duration	pkts/flow	pkt/s/flow	bytes/pkt
IP in IP		✓	0.04	2.73	2.57	0.09	456	173.1	2307	10.8	253
TCP ftp-data	20	✓	0.76	12.09	15.18	2.17	2018	118.2	525	3.8	338
TCP ftp-cntrl	21		1.55	0.74	0.23	6.50	124	38.6	16	1.1	83
TCP telnet	23	✓	1.39	4.81	1.61	4.24	803	114.3	114	1.1	90
TCP smtp	25		10.26	4.80	2.82	49.49	802	18.2	15	2.0	158
UDP dns	53		45.30	5.57	3.04	216.56	929	15.4	4	0.5	147
TCP gopher	70	✓	0.45	0.54	0.55	1.87	91	43.3	40	2.2	275
TCP http	80	✓	17.94	40.21	41.53	72.98	6717	56.5	74	2.5	278
TCP pop-v3	110		0.08	0.05	0.03	0.41	9	27.0	21	2.5	148
TCP authent	113		2.12	0.19	0.05	10.54	32	9.0	3	1.5	64
TCP nntp	119	✓	0.35	6.56	6.59	0.68	1096	176.7	627	3.6	270
UDP ntp	123		5.01	0.20	0.06	25.02	33	1.37	1.3	1.3	83
TCP netbios	139	✓	0.03	0.08	0.15	0.11	14	69.8	82	1.2	501
UDP snmp	161		1.35	0.26	0.11	6.14	43	17.9	6	0.7	115
TCP login	513	✓	0.09	0.24	0.14	0.31	41	88.1	92	1.3	156
TCP cmd	514	✓	0.01	0.13	0.07	0.06	21	49.1	316	3.7	149
TCP audio	1397	✓	0.00	2.20	2.62	0.01	367	167.9	15653	73.8	321
TCP AOL	5190	✓	0.18	0.46	0.38	0.51	77	129.8	84	0.7	223
TCP X-11		✓	0.08	0.66	0.53	0.18	111	160.6	276	3.3	217

Table 1: Flow Statistics per Protocol

percentile of 116 flows per second. The average number of established flows in the connection table is 15,500. This rate of connection requests, and the connection table size, are within the capabilities of current ATM switches.

The experiment was repeated with all packets classified for switching. Thus, a suitable connection must exist, or be established, for every packet. This simulates the purely connection-oriented approaches to IP over ATM. In this case a mean of about 420 flows/s must be established, about 4.5 times the rate of connection requests as in the previous experiment, with an average of over 40,000 entries in the flow table. This rate of connection requests, and the size of the connection table, would stress current ATM switch designs, and it is generated by a traffic stream with an average bit rate of only 36 Mbits/s. This clearly demonstrates the advantage of connectionless forwarding for packets belonging to short lived flows.

The actual value of establishing a connection for long lived flows is dependent upon the amount of work required to perform the signalling operation and the delay before the connection is established. To consider an upper bound on performance let us assume that it takes on average six signalling messages to establish a connection with zero setup delay. So the amount of work required for the signalling operation is very roughly equivalent to the amount of work required to forward 7 IP packets (6 signalling messages plus the original IP packet that must be forwarded before the connection is established).

The trace contains an average of 16,700 incoming packets per second of which, in the second experiment, about 14,100 are recognized as suitable for switching and the remaining 2,600 must be forwarded by the processor. Thus, the amount of work required to establish 92 connections per second is approximately equivalent to the work required to forward 640 packets per second. So, if connections are established for all of the packets marked as suitable for switching, 16,700 packets per second may be handled with an amount of work equivalent to that required to forward 3,240 packets per second. By establishing switched connections for long lived flows we are able to handle an upper bound of about 5 times more traffic than if all packets were forwarded. However, we have reached the limits of the signalling protocol for current generation ATM switches and are only switching a traffic stream of 36 Mbits/s.

Acknowledgement

We are grateful to K. Claffy and Hans-Werner Braun, Applied Network Research, San Diego Supercomputer Center, for making the trace available to us. The trace is available at <ftp://www.nlanr.net/Traces>.

References

- [1] K. C. Claffy, H.-W. Braun, and G. C. Polyzos "A parameterizable methodology for Internet traffic flow profiling," IEEE J. Selected Areas in Commun. 13(8), Oct. 1995, 1481-1494.