#30

ATM Experiments for Advanced Backbone Services

Michael Behringer

This paper was presented at INET'97, the Annual Conference of the Internet Society, Kuala Lumpur, Malaysia, June 1997

DANTE
Francis House
112 Hills Road
Cambridge CB2 1PQ
United Kingdom

Tel: +44 1223 302 992
Fax: +44 1223 303 005
E-mail: dante@dante.org.uk

ATM Experiments for Advanced Backbone Services

Michael Behringer

Abstract

Most production ATM backbone networks make use only of conservative ATM traffic classes such as CBR or VBR, avoiding the difficulties of more complex ATM services. In Europe theTEN-34 project aims at providing more advanced ATM services on a pan-European scale. This paper describes the experiments carried out to evaluate the applicability of advanced ATM services such as SVCs and NHRP in an international backbone environment. The tests are defined and carried out by the TERENA task force TEN (1), using the pan-European JAMES ATM backbone (2). The set of experiments described here provides a concise overview of the usability of advanced ATM features in the wide area.

1. Introduction

Up to the beginning of 1997 Internet backbones in Europe consisted mostly of 2 Mbit / s line technology. In 1996 the national research networks (NRN) in Europe with DANTE as project manager started a project - similar to the Internet II efforts in the USA - to develop a pan-European backbone with higher speeds and more advanced technology: TEN-34 (Trans-European Network Interconnect at 34 Mbit / s). As there was desperate demand for more Internet bandwidth in the short term the first phase of the TEN-34 project (3) was limited to provide a standard high-speed Internet service for the academic community. This TEN-34 network is now operational (4). The need for more advanced applications was however recognised, so the second goal of the project was to investigate into new technologies such as bandwidth on demand, to support new applications, and to make them available on the TEN-34 network.

This paper describes the experiments carried out in the framework of the TEN-34 project to examine new technology trends, mostly based on ATM, on their suitability for the provision of backbone Internet services. A task force, consisting of experts of many European countries, was established to carry out these experiments. This Task Force TEN (Trans-European Networking) (1) is organised under the framework of TERENA (5).

The experiments described here cover a wide range of aspect of ATM, from network management to signalling, as well as performance related issues and in depth examinations of ATM technology related issues such as cell delay variation. Unfortunately the results are far from encouraging. In several areas of research we came to the conclusion that too many pieces of the jigsaw are still missing. Especially the use of SVCs does not seem to be feasible over a WAN yet.

But why did we discover so many problems with ATM in the wide area, where there seem to be no problems in the local area? After all, SVCs are already successfully being used in LANs. The difference is that in the LAN bandwidth is close to "unlimited". As soon as there are bandwidth restrictions however as is the case on WANs, it becomes far more difficult to accommodate many different traffic flows. And software and hardware has not yet adapted the mechanisms needed to cope with low bandwidth. Clearly ABR was designed to solve some of these issues, and also implementations of device drivers and switch software will improve. So our conclusion is not that the more advanced ATM features will not work, but it is apparent that far more time is needed to deliver on the promises that were made on ATM.

The "Internet technology" is developing further as well, and it is not apparent yet whether either side can solve all the problems. Thus we are also looking into RSVP as an alternative model for reserving bandwidth. The state of NHRP is also under investigation. In these times of semi-religious arguments about ATM versus IP it is important to stress that we are not biased towards either technology. We have real demands, and the best technology willbe deployed.

Michael Behringer holds the position of Senior Network Engineer in DANTE. His e-mail address is Michael.Behringer@dante.org.uk

Quite a few buzzwords of ATM, such as ABR and PNNI do not appear in this paper. The reason is not that we do not consider them, but implementations of them were either not available, or we didn't have time to do experiments on them. The second phase of the TF-TEN tests, which started in May 1997, will cover these areas, and continue some of the tests described here.

2. Experiment Descriptions

In the following sections the results of our experiments are outlined. Due to space limitations it was not possible to provide extensive explanations of the results. In general experiments in which we discovered problems are covered in greater length. For more detailed explanations please see our web site (1) or get in touch with the experiment leaders directly (see acknowledgements). This list contains the experiments carried out so far:

1. TCP / UDP over high-speed and long delay
2. SVC testing
3. ARP testing
4. NHRP testing
5. Addressing Issues
6. Network Management
7. CDV Tests:
8. IP over VBR
9. RSVP
10. ATM Security

2.1. TCP/UDP over high-speed and long delay

The purpose of this experiment was to verify the validity of existing research results over ATM VPs, and to ensure that there are no new problems introduced through ATM. For this test two long-distance VPs were used and the same set of experiments carried out: One VP was 22 Mbit / s and went from Italy to Sweden (48 ms RTT), the other one was 14 Mbit / s (36,000 cells / s) and went from Spain to Norway (63 ms RTT).

The experiments carried out were TCP and UDP throughput measurements between one and more hosts. The workstations had to be patched to allow bigger windowsizes for TCP. Netperf was used in single user mode to do the tests. The result of both tests was that a utilisation of the VPs of almost the full theoretical maximum could be reached. We did not discover any unexpected interference with the ATM layer, and the results were as expected.

2.2. SVC testing

SVCs are being used successfully in local area networks already. This test was to check the feasibility of ATM signalling and switching over policed wide area links provided by the JAMES network. One problem was that the JAMES network in use did not support switching. To circumvent this, we tunnelled the switching information through CBR permanent VPs by moving the switching channel (usually VPI 0, VCI 5) into the VP provided by the network (e.g., VPI 10, VCI 5) on the switches that interfaced the CBR service. This way the signalling would be done on switches in two countries and tunnelled through the public network over a normal CBR virtual path connection.

The tests showed that switching over a policed network is highly unstable and hardly unusable for an operational service. There are a number of problems that have to be resolved before switching can be used in its intended way, between end-applications:

- Applications need to support the ATM functionality by being able to specify, for example, required bandwidth. Only prototypes are available so far.
- The TCP / IP stack must be able to pass on the QoS information between the application and network. There are prototype implementations for this, but there are no supported TCP / IP stacks where applications can request for example a 2 Mbit / s VBR SVC with certain parameters. It has been shown though that this can be done by e.g. the TCP ONIP (TCP over non-existent IP over ATM) implementation (6).
- Routers and switches must do accounting of available bandwidth. At the time of the SVC set-up the lowest bandwidth on the whole path must be determined and negotiated with the application.

These issues must be resolved and all the equipment between end- applications needs to support all those features to make SVCs work in the way that was intended. Availability of these features in fully supported software releases of workstations, routers and switches is not yet in sight.

Although this full solution might be still some years ahead, one could envisage the use of SVCs in a more constrained way to bypass some of the problems. For example it would be possible to leave the "accounting" of available bandwidth to the human user. Prototype applications of TCP implementations and of applications are available. But such a solution would first of all not meet the stability requirements needed in an operational service. And even then, there are still a number of problems in today's systems:

- We found a mismatch in VCI numbers between the sending and the receiving system, which led to a high percentage of connection set-up failures. If the sending side chooses a VCI which not in the range of accepted VCIs on the receiving side, the connection will be refused. There are commands to limit the range of VCIs, but we found them not to work reliably on some types of equipment.
- A number of workstation ATM cards defaults to using a UBR service at line speed. This means that an STM-1 card for example sends traffic on SVCs at 155 Mbit/s. If the bandwidth available on the path to the destination is lower, the result will be close to zero "goodput". Single packets can get through up to the size that fits into the number of cells permitted in the CDVT of the network. So we always got through 108 byte pings, but never 109 byte pings. Whilst this is a nice way of identifying the CDVT of the network, it is otherwise quite limiting. New drivers allow the definition of a maximum speed for SVCs, but this applies to all SVCs, which means that two concurrent VCs would compete for this bandwidth, resulting in very low "goodput". Shaping on the edge to the public network is a possibility to circumvent this problem, but it can hardly be called a good solution.
- There also seem to be several problems in the signalling protocol or the implementations of it as such, which we could not identify clearly yet. In certain cases a workstation that received a call set-up created a new SVC for the response instead of using the established SVC. Quite often the call was rejected for no apparent reason. In other cases the VCs were established but not used by the IP stack.

One way to avoid some of these problems is to do traffic shaping on a switch before the public network, which is policing the traffic. In this case it is the switch that would throw away the cells and the flow control is left to TCP. This solution works well. In our tests we noticed almost no cell losses, as there was sufficient buffering on the switches. Up to 10 concurrent TCP connections were tested over CBR VCs with 2 Mbit/s, and we could observe a fair sharing of bandwidth, as can be expected from using TCP. This solution however works only for TCP, and one UDP connection can break the whole system. We did not look into packet discard mechanisms here.

We also investigated in the set-up and tear-down times for international SVCs by sending a number of pings, in which the first one is longer than the subsequent ones, due to the set-up of the SVC (if the calls succeed). The distribution of set-up times was inconsistent and did not meet our expectations. More research on this is currently being undertaken.

In summary we concluded that SVCs in the way they were intended to be used, between end applications, will still not be feasible for the next two to three years. Even on a very limited scale, where the number of SVCs on a given VP never exceeds one and all the parameters are hard coded, the current implementations of drivers, applications and router and switch software seem to be far to unstable to provide any operational service on top.

2.3. ARP testing

The goal of this experiment is to test the usage of an ATM ARP server in the wide are network. As ARP servers have been in use for some time in local area networks, the focus of this experiment was more on the timing of the ARP requests and responses over a wide area network.

The set-up in use was one ATM ARP server in Austria, which was used by the other participants of the SVC experiment to resolve IP addresses to the NSAP addresses we used. As expected we did not discover any major problems and the ARP server was used in a normal way to support the other experiments where needed.

More experiments with several ARP servers and logical IP subnetworks will be carried out, to test the scalability. However, no major problems are expected.

2.4. NHRP testing

The Next Hop Routing Protocol is way of finding an optimal shortcut through large public (Non-Broadcast Multi-Access) networks like X.25, SMDS and ATM. It bypasses the normal IP routing paradigm by being to establish direct connections to systems on a different IP-subnetwork. In this way it is possible to support interworking of thousands of systems belonging to different organisations on a common infrastructure in an efficient way.

The experiment we carried out first involved two hosts in Norway and one in Spain. We used Cisco routers and the NHRP implementation of IOS version 11.0. This set-up worked as expected, and there were no major problems, apart from the usual SVC problems (see section 2.2.), and even those were partly tackled through the shaping done on the router. We experienced some routing instability problems on our router with cache invalidations probably due too low software levels.

We are currently setting up a test environment with more than one IP hop, to test the scalability, and more of the parameters such as the number of packets received before the shortcut is activated. For a big IP network the question remains whether this solution will scale with the standard CBR or VBR services: Due to the large number of potential connections over the network and the fact that bandwidth is limited this would lead to either very small VCs or only very few of them. Before services such as ABR are available the n-to-n problem might still prevent NHRP to be deployed widely in a backbone environment.

2.5. Addressing Issues

The goal of the Addressing investigations is first to identify the addressing schemes in use in Europe's research environment and in the public ATM networks, and to develop experience on address conversion should this become necessary.

A survey amongst the national research networks (NRN) in Europe showed that most NRNs are using or are planning to use DCC format NSAP addressing for their ATM networks (7, 8, 9, 10). The public ATM network operators have not come to an agreement on which addressing scheme to use commonly, and it does not look likely that there will be an agreement. This means that an address translation function will be needed to make ATM addressing work globally. However, this requirement has not yet been acknowledged by the ATM operators, as they are still fighting for one common addressing scheme.

Supported address translation mechanisms between E.164 and NSAP addressing had not been available at the time of writing this paper, and only beta versions were becoming available. Therefore no practical experience in address translation could be gained yet.

2.6. Network Management

ATM networks will require different and more sophisticated ways of network management than traditional IP networks, as the functionality increases with e.g. signalling. The goal of this experiment series is to evaluate the different available network management options and to trial them on our international ATM network.

The traditional network management protocol used in the Internet is SNMP, which we concentrated on in the first phase of the tests. The switches in use were polled from one central NM platform in Belgium both through the ATM connections (in band) and through the production Internet (out of band). We used standard MIBs where available, and also manufacturer specific MIBs. In none of the cases we experienced any problems. The results of the polling were made available on a web page, so that all project participants could check the status of the network.

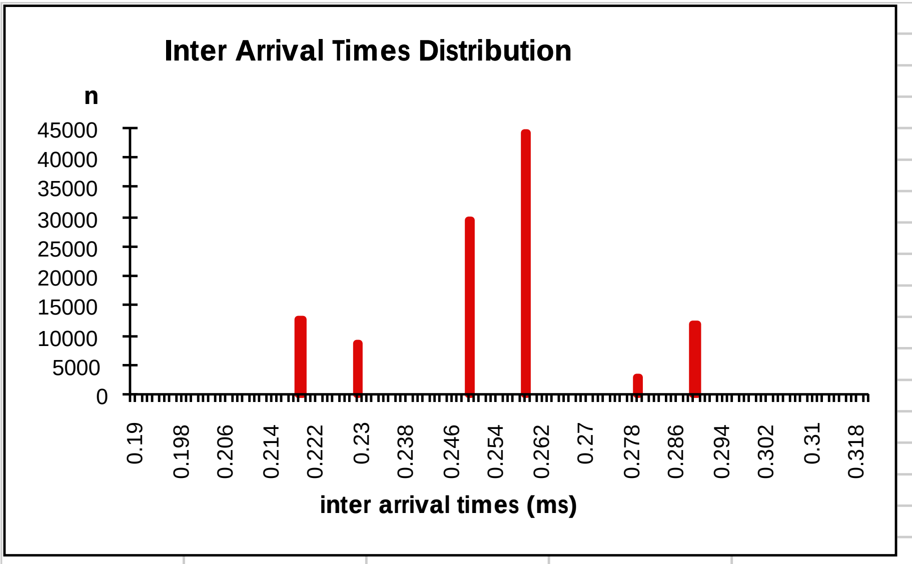
In the next phase of the project we envisage trialling different NM protocols such as the X-user interfaces. The JAMES network plans to support this feature for its users. At the time of writing this paper no experience on this protocol could be reported.

2.7. CDV Tests

ATM CBR services promise to deliver a constant data stream with very low jitter, i.e., small delay variation over the network. This feature is essential for the support of applications such as PDH circuit emulation and transfer of certain kinds of video data. This experiment aims at evaluating the CBR services on their quality of service in this respect.

To evaluate the cell delay variation of a single stream of data over an ATM VC, we established an international VC with several measurement points on the path, where the delay variation could be measured with the help of ATM analysers. The longest path in test involved five such checkpoints. Over this path we sent an (almost) 0 CDVT cell stream and measured the arrival times for each cell on each checkpoint. The distribution of cell interarrival times was then analysed.

The results were far from what could be expected from a CBR service. Whilst at the beginning of the VC the cell interarrival time varied only by 3-10 µs, with each switch that was passed the cell interarrival time variation increased, and on the receiving side interarrival time variations of up to 130 µs could be observed. At the time of writing this paper no explanation for this behaviour could be found yet. Especially since the used service was a CBR service this finding is at least concerning.



This figure shows the inter arrival times measured at the last checkpoint in the path, where a significant variance can be seen. When comparing the measurements on different checkpoints, one can see that the further away from the source, the wider the spread of the interarrival times, and the longer the average interarrival time.

These results suggest that in large ATM networks re-shaping might become necessary after a number of switches, as each switch adds to the variation in interarrival times. We have some experience in using VPs with CDVT of around 300 µs. This still seems to be a reasonable value for current ATM networks, however, it might need to be reviewed with significantly growing ATM networks.

2.8. IP over VBR

The goal of the VBR experiment is to assess the technical advantages of using VBR for IP traffic, as opposed to CBR. At first glance it looks like VBR provides the same sort of flexibility as Frame Relay does. However, on a VBR service the peak cell rate (PCR) can only be used for very short periods of time over the guaranteed average rate, the SCR (sustainable cell rate). The burst tolerance (BT) specifies the duration of the peak, and is usually in the range of a few cells. After a peak the transmission has to slow down for some time, to reach on average the SCR again. Thus a VBR service does not provide "additional bandwidth" on top of the SCR for IP traffic.

The main question of the experiment is whether there is any benefit at all for IP traffic on a VBR service. Experiments on VBR services were carried out in the Netherlands (11) and in Switzerland, in both cases locally. The main results from the Dutch tests were:

- The IP throughput is mainly determined by the SCR. It will not go significantly beyond the SCR, and certainly not to the PCR.
- If the BT on the router is higher than the burst tolerance on the ATM network, the "goodput", i.e., the usable IP throughput is close to zero. This is not surprising, as the policing of the ATM network would discard cells that do not correspond to the traffic contract, and as any cell loss leads to severe IP packet losses.
- Higher BT (within the traffic contract) leads to slightly higher throughput. This is also not surprising.
- Measurements with PCR = 2 * SCR showed slightly lower throughput than with PCR = SCR. This is surprising, since an increase in PCR should increase the performance, rather than decrease it. Up to the time of writing this paper no explanation for this behaviour could be found.

For more information on these results see (11). A confirmation of these results over an international VP was being carried out at the time of writing this paper. The results shown here demonstrate that VBR services can be used, but they do not offer any advantages over CBR services for IP traffic, except that on the ATM network itself it is cheaper to provide VBR services. Configurations should always be with PCR=SCR, and the BT should be as big as possible.

2.9. RSVP

The goal of this experiment is to evaluate the RSVP protocol on its suitability for bandwidth reservation and to check the interworking with the ATM layer. For this test we established first a connection over one logical IP subnetwork (LIS), then we traversed two LISes with a router in between. So far only lab tests have been conducted.

In both cases the ATM VCs were established as expected and the performance of the set-up was stable. These tests are currently being re-evaluated in the wide area. The major problem is that the software used for the tests is not supported, which currently makes the introduction of RSVP in the production network impossible.

RSVP multicast has not been considered in this phase of the project.

2.10. ATM Security

With ATM networks becoming more sophisticated there is an increasing threat that new capabilities could be used in other ways than originally intended, for example by faking signalling messages. The research carried out in this context focuses only on security related to the ATM functionality, such as switching and other interactions with the public ATM network. Other aspects such as sniffing and security on higher protocol layers are not considered here.

So far the work carried out is only theoretical. A threat analysis and more related information can be found on the TF-TEN homepage (1). Practical experiments are expected to start in May 1997, and results will be available on the above mentioned site.

3. Summary and further work

The work carried out in this framework shows that most of the advanced features of ATM and the new IP protocols are not yet in a state where they can be used safely for operational services. The problem seems to be in most cases that the development of hard- and software is not mature enough. The results of the experiments do however show how to make best use of the existing services (CBR, VBR), and give a good insight into the problems that arise with new technologies.

All the experiments listed here are part of the first phase of the TEN-34 testing programme, which was not finished at the time of writing this paper. More work is clearly needed to fully understand the capabilities of ATM networks and of comparable IP services. In some of the areas described above new questions arose during the tests.

There are also a number of technologies which were not yet examined in phase one. Phase two of the project, starting in May 1997, will also investigate into other technologies, such as ATM routing and new traffic classes such as ABR. The focus of the tests carried out here is to make experimental services available on the production TEN-34 network. Although the more interesting features of ATM seem to be not in a state yet for an operational service, we will keep on following the developments in ATM and IP related activities. The latest information on our experiments can always be found on the TF-TEN home page (1).

Acknowledgements

The work presented here is the work of the TERENA Task Force TEN (1). Whilst the author is chairing this task force, the work was carried out by the members of the group. This being a European project, a large number of people were involved, far too many to list them here. The leaders of the experiments are listed here on behalf of all the other researchers involved in this work.

Experiment Leaders and their e-mail addresses

TCP / UDP performance
Mauro Campanella CAMPANELLA@mi.infn.it
Tiziana Ferrari Tiziana.Ferrari@cnaf.infn.it
SVC testing
Christoph Graf Christoph.Graf@dante.org.uk
ARP testing
Ramin Najmabadi Kia
Ramin.Najmabadi@HELIOS.IIHE.AC.BE
Simon Leinen simon@switch.ch
NHRP testing
Olav Kvittem Olav.Kvittem@uninett.no
Addressing Issues
Kevin Meynell K.Meynell@terena.nl
Network Management
Zlatica Cekro cekro@HELIOS.iihe.ac.be
CDV Tests
Victor Reijs Victor.Reijs@SURFNET.NL
IP over VBR
Olivier Martin omartin@DXCOMS.CERN.CH
RSVP
Olav Kvittem Olav.Kvittem@uninett.no
ATM Security
Paulo Neves pneves@rccn.net

References

1. TF-TEN stands for Task Force on Trans-European Networking; see http: / / www.dante.net / ten-34 / tf-ten /

2. JAMES : See http://www.labs.bt.com/profsoc/james/
3. TEN-34 : Project to establish a pan-European high-speed research backbone (Trans-European Networking on 34 Mbit/s). See http://www.dante.net/ten-34/
4. Behringer, Michael: "TEN-34 and JAMES: Technical Plans"; Proceedings JENC7, Paper 131, 1996.
5. TERENA : For more information see http://www.terena.nl/
6. Afifi, Hossam: : "TCP over non-existent IP over ATM Networks", Proceedings JENC7, Paper 153, 1996. http://www.terena.nl/
7. Dick Gillman, George Howat, Jim Jackson, Baoyu Wang : "Interim Private ATM Addresses for the UK Academic Community"; March 1996; http://www.ed.ac.uk/~george/ukac-addr.html
8. Preliminary SURFnet (NL) Addressing Plan: http://www.nic.surfnet.nl/surfnet/projects/atm/nednsap.htm
9. DFN Addressing Plan (DE): http://www.dfn.de/dfn/dfn-wid/nsap-inst.html
10. Supernett (NO) Preliminary Addressing Plan: http://www.uninett.no/info/nett/supernett/atm-adresser.html
11. Reijs, Victor: "VBR and IP"; http://www.nic.surfnet.nl/surfnet/persons/reijs/sn4/pcr.htm ; 1996.