THE AUSTRALIAN NATIONAL UNIVERSITY

TR-CS-98-04

ACSys/RDN Experiences with Telstra's Experimental Broadband Network, Second Progress Report

M. Wilson and K. Yap

April 1998

Joint Computer Science Technical Report Series

Department of Computer Science
Faculty of Engineering and Information Technology

Computer Sciences Laboratory
Research School of Information Sciences and Engineering

This technical report series is published jointly by the Department of Computer Science, Faculty of Engineering and Information Technology, and the Computer Sciences Laboratory, Research School of Information Sciences and Engineering, The Australian National University.

Please direct correspondence regarding this series to:

Technical Reports
Department of Computer Science
Faculty of Engineering and Information Technology
The Australian National University
Canberra ACT 0200
Australia

or send email to:

`Technical.Reports@cs.anu.edu.au`

A list of technical reports, including some abstracts and copies of some full reports may be found at:

`http://cs.anu.edu.au/techreports/`

Recent reports in this series:

TR-CS-98-03 M. D. Wilson, S. R. Taylor, M. Rezny, M. Buchhorn, and A. L. Wendelborn. *ACSys/RDN experiences with Telstra's experimental broadband network, first progress report*. April 1998.

TR-CS-98-02 M. Manzur Murshed and Richard P. Brent. *Adaptive AT^2 optimal algorithms on reconfigurable meshes*. March 1998.

TR-CS-98-01 Scott Milton. *Thread migration in distributed memory multicomputers*. February 1998.

TR-CS-97-21 Ole Møller Nielsen and Markus Hegland. *A scalable parallel 2D wavelet transform algorithm*. December 1997.

TR-CS-97-20 M. Hegland, S. Roberts, and I. Altas. *Finite element thin plate splines for surface fitting*. November 1997.

TR-CS-97-19 Xun Qu, Jeffrey Xu Yu, and Richard P. Brent. *Implementation of a portable-IP system for mobile TCP/IP*. November 1997.

DHPC Technical Report DHPC-023
ANU Technical Report TR-CS-98-04

ACSys/RDN Experiences with Telstra's Experimental Broadband Network

Second Progress Report

M. Wilson and K. Yap

Distributed High Performance Computing Project
Cooperative Research Centre for Advanced Computational Systems,
Research Data Networks Cooperative Research Centre

Originally Released June 1997

1. Abstract

This report addresses three issues relevant to DHPC network infrastructure requirements:.

- Technically, what we can currently achieve with the EBN?
- What are desirable qualities in future ATM networks?
- What further development of the EBN would we like to see for other applications?

In writing this report, we will briefly describe the EBN ATM experience of ACSys' Distributed High Performance Computing Project and Digital Media Libraries project, explain some of the technology and networking issues we have dealt with and summarise what we have learned. We have also included some background material on the EBN and ATM networks in general.

2. Introduction

This report and its predecessor[17] describe work which was undertaken in 1997 and the second half of 1996. Initially intended as progress reports to Telstra, the provider of our experimental ATM [10,11,12,13,14,15,16] testbed, they are now being published as technical reports because they describe some of the important technical hurdles we overcame, the lessons we learned, and provide a convenient reference for discoveries which form a basis for much other work which has been undertaken by the DHPC [7].

3. Introduction to the EBN and ATM Terminology

The EBN [1] currently supports two switching services - Permanent Virtual Circuits (PVCs) and Permanent Virtual Paths (PVPs). To try to avoid confusion, we refer to these two services as EBN PVPs and EBN PVCs, to distinguish them from circuit connections we configure ourselves – particularly in sections 4 and 5.4, where we discuss networks of PVCs and SVCs to be switched together within EBN PVPs.

Telstra doesn't plan to provide a Switched Virtual Circuit (SVC) service on the EBN.

The following are described in the context of the EBN, but can be assumed to apply to ATM networks generally.

3.1 Permanent Virtual Circuits

A Permanent Virtual Circuit, (also called a Virtual Circuit Connection), is a dedicated unidirectional point-to-point circuit between two ATM endpoints. A PVC is identified at each endpoint by a pair of numbers - a Virtual Path Identifier (VPI) and Virtual Circuit Identifier (VCI). ATM Cells entering the ATM network are tagged with these identifiers, which are then used by intermediate network switches to route the cells to their destination. Switches within the ATM network can change these identifiers as the cells are switched, so there is no requirement that the endpoints of a PVC have the same VPI/VCI pair. These Virtual Circuits are Permanent in the sense that an operator must configure them manually; they are not allocated on demand. Because PVCs are unidirectional, useful data connections will make use of a pair of PVCs, one in each direction. In this report, we will use the acronym "PVC" as a synonym for such a pair of PVCs. Using the EBN's PVC service, bandwidth resource is allocated permanently on an individual PVC by PVC basis.

3.2 Switched Virtual Circuits

Like Permanent Virtual Circuits, Switched Virtual Circuits (SVCs) provide point to point connections between two endpoints. Unlike PVCs, SVCs are established on demand using a standard signalling protocol. Because SVCs are created automatically by the network on behalf of applications, the manual administrative overhead associated with an SVC-based configuration is potentially much lower than a network of PVCs. The signalling protocol used to establish SVCs also accepts Quality of Service parameters, including the required bandwidth in each direction. In short, SVCs are more flexible than PVCs.

SVCs are not provided directly by the EBN, but we could potentially establish a private SVC service using the signalling capabilities of our own switches – this is also discussed in section 0.

3.3 Permanent Virtual Paths

Permanent Virtual Paths, also referred to as Virtual Paths or Virtual Path Connections, are a mechanism for conveniently switching a group of Virtual Circuits (PVCs or SVCs) between common endpoints. PVPs are distinguished by the VPI identifier only; cells from an ATM site with a given VPI are switched through the network together, irrespective of their VCI numbers which are left unchanged. The Distributed High Performance Computing Project (DHPC) has found the EBN PVP service useful because it allows us to manage our own PVCs without outside assistance from the service provider. When using the EBN's PVP service, a portion of an EBN connection's bandwidth is allocated permanently to the Virtual Path as a whole, leaving bandwidth allocation between individual PVCs to the customer's or researcher's network administrator.

Some commercial ATM switches, including the FORE ASX-1000, the IBM 8260 and Digital GIGASwitch/ATM can utilise a Permanent Virtual Path to implement SVCs. This is called Virtual Path Tunnelling. We have attempted to implement this capability between Adelaide and Canberra. Unfortunately we have experienced some technical problems, but we are receiving assistance from Digital.

3.3.1 Classical-IP

We are able to run the Internet Protocols [2, 3] on our local ATM LANs using a standard called Classical-IP [4, 5]. Classical-IP can make use of SVCs to provide VC connections between hosts as they are required, and to retire connections automatically as they become inactive. Classical-IP also supports PVCs, which is how we currently communicate over the EBN. Classical-IP is a simple but well-defined standard, which is widely supported.

Another implementation of IP over ATM is ATM LAN Emulation, or LANE [6]. While LANE is a more comprehensive standard than Classical-IP (LANE provides a broadcast facility, for example, where Classical-IP doesn't), we have not changed our implementation because Classical-IP has satisfied our requirements so far. LANE is certainly worth investigating, because it opens up possibilities for IP switching and Virtual LANS across the EBN. We'll be experimenting with changing the Canberra ATM LAN over to LAN Emulation in the near future.

3.3.2 Flow Control and Traffic Management

Flow control is an important issue in ATM networks that are used to transmit computer data. ATM is not a store and forward protocol – it does not guarantee delivery so that the implementation of ATM switches can be made simple and fast. Traffic can be lost in an ATM network if a host transmits data too quickly for any intermediate network link. The switch feeding the congested link may try to buffer the data, depending on the kind of service provided, but congestion will ultimately lead to data loss as the switch drops excess cells. In addition, the ATM switch does not inform the sending host that the data has been dropped - this is left to a higher layer protocol. On a TCP connection, cell loss can be observed as poor throughput. The EBN switches (supplied by Alcatel) were originally equipped with small cell buffers, which caused EBN users to observe poor performance. The buffers on the EBN ingress ports were replaced in late 1996, which significantly improved network performance.

In order to manage traffic, a number of ATM service types [11, 12, 19] were conceived by the ATM Forum , an influential U.S. consortium of ATM vendors:

- A Constant Bit Rate (CBR) service provides what you might expect - customers are allocated a maximum number of cells per second, defined by a minimum inter-cell spacing. The customer must not exceed the contract, or data is lost. This service is ideal for voice, uncompressed video communication and other timing-sensitive applications. Internally, the EBN is effectively a CBR network.

- A Variable Bit Rate service (VBR) is described by an sustained cell rate, combined with a burst allowance (which the customer pays for). This was originally designed for applications like compressed video, but the compressed video experts think VBR is better suited to data traffic. Unfortunately, VBR doesn't suit data communications well either. Externally, the EBN provides a service more like VBR than CBR, as a degree of burstiness is tolerated in incoming traffic. The ATM Forum actually defines two VBR services - one that addresses timing relationships in ATM traffic, and one that doesn't.

- What best suits bursty computer data communication is an Available Bit Rate (ABR) service, which allocates bandwidth on demand from that which is available. Each ABR connection using a service can use anywhere between nothing and the full available capacity, so connections need not have bandwidth allocated to them in advance. Because a number of applications may simultaneously be using the common bandwidth pool, ABR relies on a flow control mechanism to manage bandwidth allocation and to limit cell loss.

The ATM forum has defined a number of flow control protocols, although only one of these (Explicit Forward Congestion Indicator, or EFCI) is currently in widespread use [18, 10]. Some vendors have implemented their own proprietary flow control protocols, which they claim are more effective than EFCI, but these do not interoperate[10].

EFCI is an end-to-end flow control mechanism; intermediate switches need not necessarily assist in managing congestion, which ultimately limits EFCI's effectiveness. EFCI can be implemented entirely in ATM host adaptors - requiring no special support from the network. When using EFCI, a host transmits Resource Management (RM) cells on each Virtual Circuit. RM cells are then returned to the sender by the host at the far end of the VC. If resource management cells do not return to the sending host, the host reduces the rate at which it transmits data, assuming that cell loss (congestion) has occurred at an intermediate switch. Switches may also set a Congestion Indication (CI) flag in a cell. A host may then reduce its transmission rate when it receives an ATM cell with a CI flag. EFCI is described as a *rate based* flow control mechanism, because cell rate is managed in response to indications of downstream congestion.

In contrast to EFCI, other (proprietary) ATM flow control mechanisms that we have encountered are described as "credit based". In a credit based system, downstream switches pass buffer information, or “credits”, to upstream switches and hosts, indicating what buffer space is available to a Virtual Circuit. The upstream device will then transmit only enough cells to fill the available buffer space. Proponents of credit based systems claim that cell loss is eliminated, while still providing rapid access to bandwidth for bursty data. It is generally agreed that credit based systems are superior to rate based systems for ATM data networks, and it is expected that credit based standards will supersede rate based standards in due course. Credit-based systems require hardware support in each ATM switch - increasing the cost of switches, but providing more intelligent flow control. Currently, all available credit based implementations are proprietary.

Unfortunately, because we are working in a multi-vendor environment, EFCI is what we have to work with across the EBN - in particular, between Adelaide and Canberra. Although the EBN is perhaps most accurately described as a CBR or VBR network, our DHPC workstations are using the EBN as an ABR network. This works to the extent that flow control takes care of managing cell rates and minimising data loss.

We have also recently gathered evidence[1], which suggests that TCP's own end-to-end rate based flow control mechanism may be operating between our hosts [9, 2].

We have found that whichever flow control mechanism is operating, be it EFCI or TCP, does not manage bursty traffic well on the link between Adelaide and Canberra, possibly because of the long round-trip time. In our experience, it also does not efficiently coordinate a shared link, as flow control is only managed on an individual connection end-to-end basis

4. EBN Experiences

We have experimented with three different EBN configurations between Adelaide and Canberra. These are described briefly in this section, along with the performance characteristics we observed, and the limitations we encountered.

Because we are unable to use EBN SVCs, our configurations have been implemented using PVCs, switched within a single EBN Permanent Virtual Path. Configuring individual PVCs is routine, however it is a manual task which involves logging onto each of our switches and creating the local PVC mappings by hand.

4.1 Our Initial Configuration

Our first network configuration used a Digital Network Integration Server (DECNIS router) to provide the Canberra connection to the EBN. We connected pairs of PVCs from each of the Adelaide hosts across the EBN to the E3 port on the router, and pairs of PVCs from each of the Canberra hosts across the ACSys ATM LAN to a Multimode Fibre port on the router. The router was then configured to route packets between the Canberra and Adelaide PVCs (and consequently between the Canberra and Adelaide hosts).

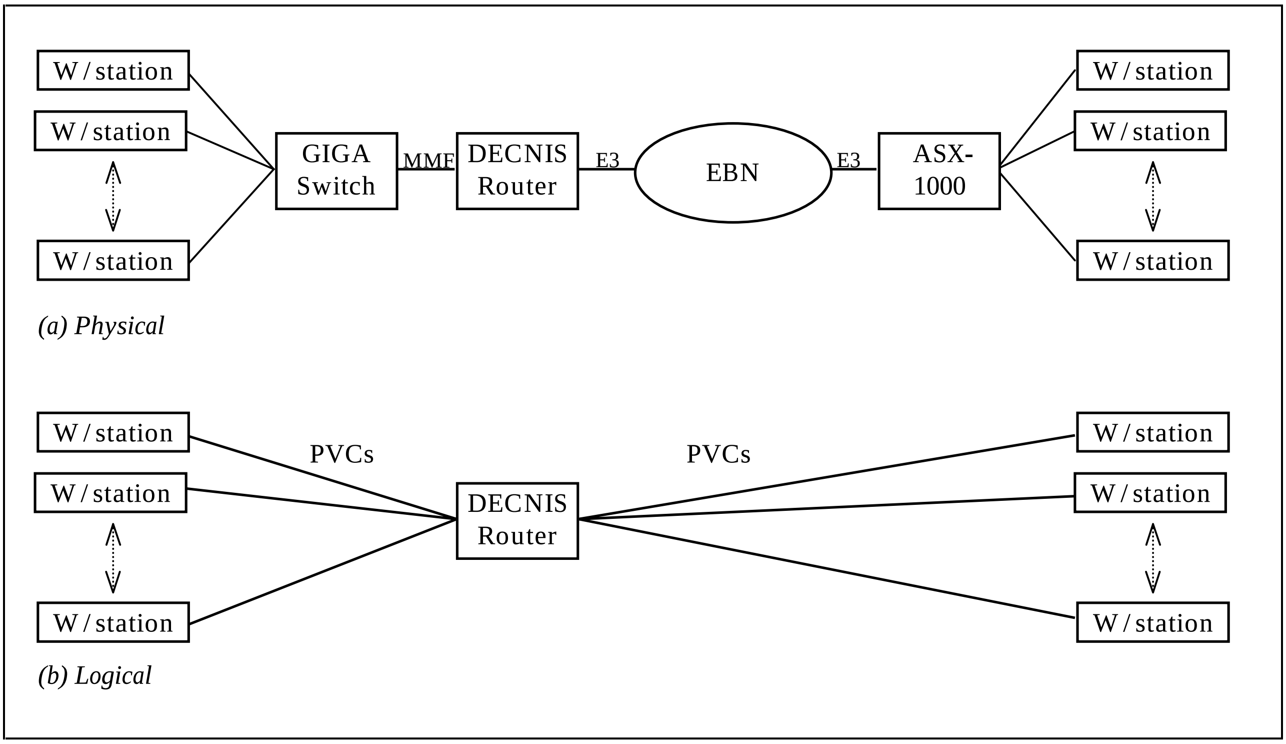

Figure 4-1: Our Initial Configuration

This configuration was moderately successful. After some performance tuning exercises on our workstations, described in an earlier report [17], we were able to utilise over 95% of the available EBN

---

[1] This theory was confirmed by tests performed By Jesudas Matthew, a PhD student working at the University of Adelaide DHPC site.

bandwidth for sustained transfers, for bursty data and even in situations where two connections were competing for available bandwidth. For more than two simultaneous connections, we saw deteriorating TCP performance. This setup was very manageable. While the router was initially a difficult piece of equipment to come to terms with, once we were familiar with it we found that changing the configuration was trivial. We were also able to arrange useful routing functions, moving switched Ethernet packets from an Ethernet-based Sun workstation over the EBN to hosts in Adelaide so that we could teleconference between Ethernet-connected hosts, while still making use of the EBN.

4.2 Our Second Configuration

The use of a router prevented us from using raw ATM (non-TCP/IP) connections across the EBN. Although we have not yet used this capability, the expectation is that we might some day be able to. The search for better network performance caused us to replace the DECNIS router with another GIGASwitch, equipped with an E3 module.

Because we could not use SVCs, we chose to configure a complete mesh of PVC pairs between the two sites - each host in Adelaide was supplied with a dedicated ABR connection to each host in Canberra, and vice versa. Creating this mesh was a non-trivial task, which involved logging into 3 intermediate switches and configuring PVCs by hand. We simplified the task by writing some simple utilities, which translated a high-level description of the PVC mesh into lists of commands in the native command language of each of our switches. While we still had to manage the switches individually, configuring and deconfiguring the PVC mesh was now a cut-and-paste process. These utilities reduced the likelihood of error when designing and configuring the PVC mesh. Each DHPC workstation was also equipped with a file describing its PVC connections which could be read at boot time, or on demand, to (de)configure the local ATM interface and routing tables.

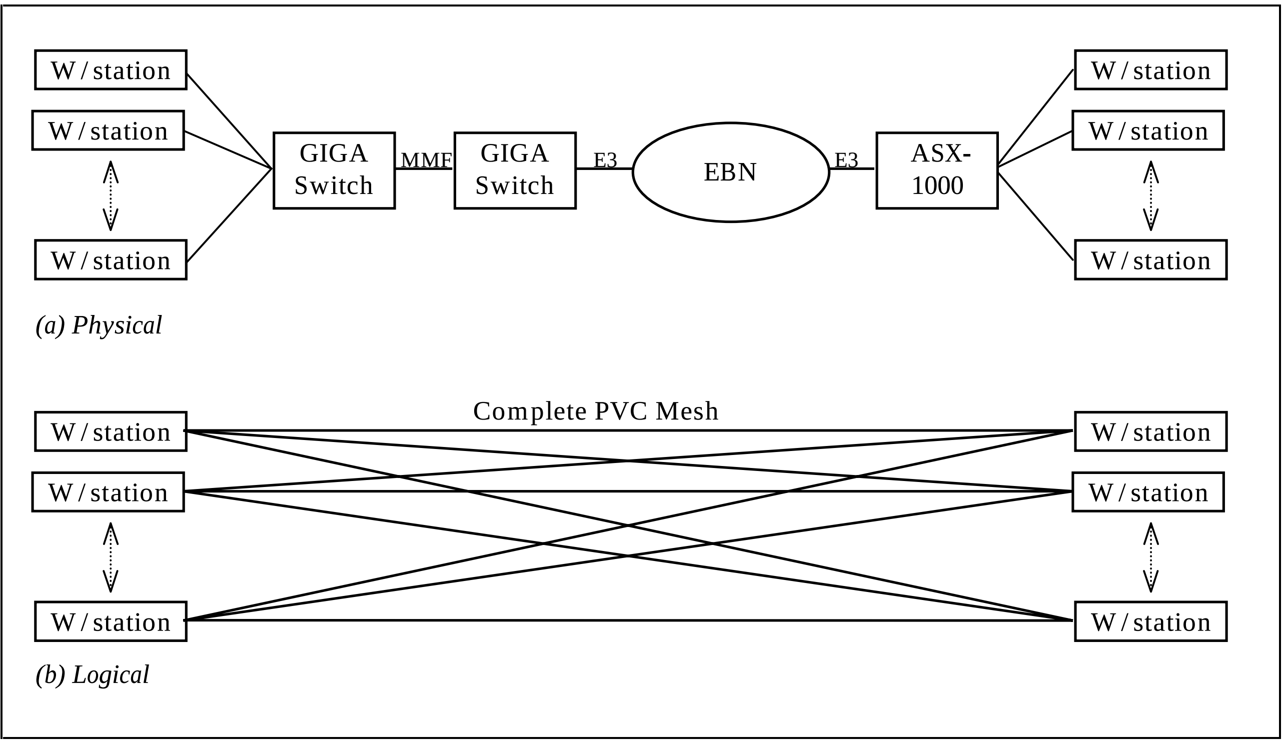

Figure 4-2: Our Second Configuration

This second configuration did not perform as well as the first. We achieved good sustained TCP throughput on a single connection, but bursty traffic performed badly. Multiple simultaneous connections also delivered diminished performance. By tuning the TCP acknowledgment window size on our workstations, we could improve the throughput of bursty traffic over TCP connections only. We have since ordered cell buffer expansion kits for our switches' line cards, and we hope this will improve the situation. We also hope that FLOWMaster, Digital's proprietary credit based flow control mechanism, will help manage bursty data and resolve bandwidth contention once we have solved some technical problems.

4.3 Our Third Configuration

The setup discussed in this section is the configuration we are currently using. We have two machines configured as routers with a bidirectional pair of EBN PVCs between Adelaide and Canberra. Other EBN PVCs establish connections to individual workstations in Brisbane and Melbourne. This configuration grew out of necessity; we wanted to establish connections to multiple sites, and our equipment at the time did not support multiple VP connections. We were not keen to establish a network of EBN PVCs between Adelaide and Canberra as this would have been a burden to maintain, a drain on Telstra's resources, and would have also partitioned our available bandwidth between individual PVCs. We are now considering using routers instead of workstations, for performance and reliability reasons.

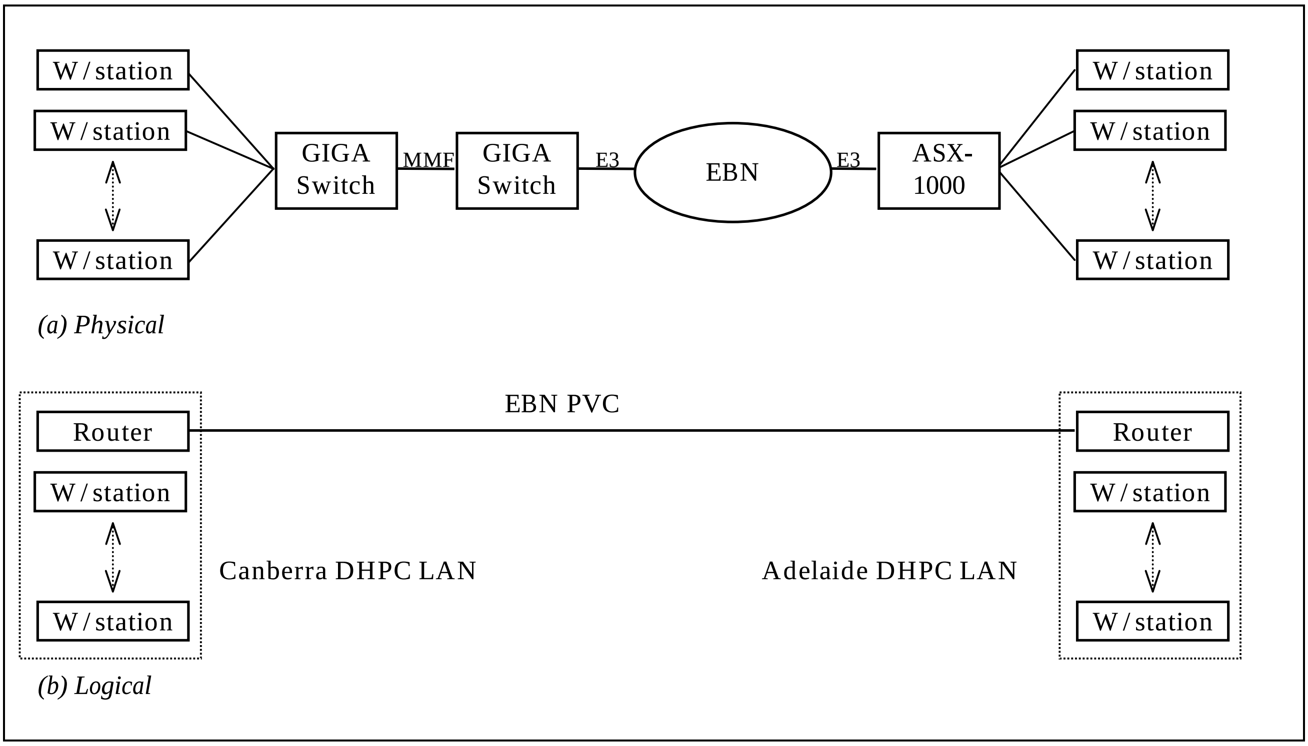

Figure 4-3: Our Third and Current Configuration

This configuration reflects the most straightforward use of ATM PVCs for data networking. It delivers good performance all round, and is straightforward to manage. Using simple throughput tests between the routers, we can demonstrate that no significant data loss is occurring over the EBN for any IP traffic. The down side is the extra processing load on the two router hosts and the restriction to IP traffic that using routers imposes. This second limitation can be overcome when necessary by creating special-purpose PVCs that switch ATM cells past the two routers. Upgrading to purpose-built routers may eventually eliminate the first problem. Tests using this configuration gave us the best indication that TCP was operating as a flow control mechanism rather than EFCI, as we previously suspected. Measured UDP streams were established between the two routers, and it was observed that beyond the rate threshold imposed by the EBN PVC, data loss was occurring with UDP that was not occurring with TCP.

The Canberra ACSys site also currently has EBN connections to the Monash ANSPAG site at Clayton and the DSTC site at the University of Queensland. These other two connections are simpler – they consist of a PVC between the router in Canberra and an individual workstation at the other site. A third PVC is used by the Digital Media Libraries group at CSIRO for a PVC connection between a pair of hosts located at Canberra and the CSIRO site at Macquarie. This last connection is also discussed in sections 4.4 and 5.4.1.

4.4 Recent Progress - Digital Media Libraries Project

A Digital Equipment Corporation GIGASwitch/ATM was connected to the Macquarie EBN site in February 1997. Initially, we encountered problems with the connection which were traced to a faulty network module on the GIGASwitch. This problem was solved by early May. We then ran into problems with a PVC between Macquarie and the CSIRO Centre for Mathematics and Information Science in Canberra. These were not solved in time, unfortunately, to run the FRANK (Film Researchers Archive and Navigation Kit) demonstration at the Research Data Networks stand at the Australian Telecommunications Users Group convention in May.

Once the PVC was working correctly, we succeeded in running FRANK from a video server in Canberra to a video client in Sydney. The control streams, which are low bandwidth, were run across the same network connection as the video stream for simplicity.

The FRANK demonstrator uses Classical-IP to transport an MPEG-1 stream at 1.5 Megabits / second between a video server and client. The stream is a constant rate so we are not affected by traffic policing because the EBN PVC is allocated a higher rate. Subjectively, the low latency of our EBN connection is quite remarkable and is, in fact, lower than that of our LAN. We believe that it's not that Ethernet has higher latency but that the TCP / IP stack on our Suns has more buffering.

At the moment, we cannot run more than one MPEG-1 stream at a time, because we only have one workstation connected to the EBN. Changing to another workstation requires changing the PVC table in the GIGASwitch and changing the setup parameters of the FORE SBA-200 interface cards at both ends. When we connect the GIGASwitch to a router, we will be able to use the EBN from any workstation on the 100 Mbit / second Ethernet LAN.

We have not connected to CMIS Melbourne for trials, but SVCs would be very desirable for this.

5. Limitations In The EBN And Our ATM Equipment

5.1 Equipment Limitations

We are currently trying to resolve several technical problems in our ATM equipment. These include problems with support for multiple Virtual Paths, Flow Control and SVC signalling. The first relates only to the Digital ATM products, while the second and third seem to be a result of an incompatibility between the Digital and FORE ATM switches. We have just received some system upgrades from Digital which promise to solve the problem with multiple Virtual Path Connections. We still need to investigate the second problem further.

Until recently, there was a serious limitation in the Digital ATM product range – their switches and routers only supported a single Virtual Path Connection, specifically VPI=0. Because we were using EBN Permanent Virtual Paths, this limited us to a single connection between Canberra and one other site, typically the DHPC site in Adelaide. The ASX-1000 used in Adelaide never had a single–VP limit, so from Adelaide we were always able to establish a VP to Canberra and to other EBN sites simultaneously. As an experiment, we have indirectly established a PVC between the Canberra DHPC site and the ANSPAG site in Clayton by switching the PVC through our VP to Adelaide, and then back across the EBN via a second Virtual Path to Melbourne. This was a terrible way to connect multiple sites, as the network transit times were very long (the ATM cells go through Melbourne twice!) leading to poorer performance than we would have achieved with a direct connection.

We have also had technical difficulties with FLOWMaster, Digital's proprietary flow control protocol. These at first glance appeared to be incompatibilities with the ASX-1000 in Adelaide. We are now trying to fix this, in the interests of minimising cell loss between the GIGASwitches at the Canberra EBN site. When we replaced the DECNIS router with a GIGASwitch, we also installed a major operating system upgrade on our Alpha workstations. Consequently, when FLOWMaster began to malfunction we couldn't determine which of the two upgrades was at fault. We may be re-testing the DECNIS router in the future to take more performance measurements and we are also receiving help from Digital, so we hope to solve this problem in due course.

We are also encountering problems using SVCs via VP tunnelling between Canberra and Adelaide. These appear to also be caused by a compatibility problem of some kind, and we are receiving assistance from Digital to try to solve these.

5.2 Bandwidth Limitations

It's worth noting that our E3 EBN connections provide us with around 27 Megabits per second useable capacity - not 34 Megabits per second. This is a design decision rather than a limitation in any of the equipment used to implement the EBN. The Alcatel switches are designed to reserve 20% spare capacity, so we only have access to 80% of the 80,000 cells/second that an E3 connection affords us, or 64,000 cells per second (~27Mb/s). If Telstra supply a 155Mb/s connection, then presumably the same 80% limitation will apply, giving us access to around 124Mb/s.

5.3 Management Limitations

There is a lot of interest in finding ways to manage network resources dynamically, to respond to application requirements. Although this is a difficult problem, we don't believe it's impossible.

The Internet Protocol suit doesn't include a standard facility for bandwidth reservation and neither does Classical-IP. Because of this limitation, any bandwidth management we implement will necessarily occur through a non-standard mechanism. It is possible that IPv6 offers us some solutions in this area which may be worth investigating. However, the lack of a standard interface to manage bandwidth at the application level is a fundamental problem.

Although ATM SVCs support bandwidth negotiation, our Classical IP implementations don't use it, treating all connections as ABR. SVCs established by Classical-IP specifically use all of the bandwidth they can get. The Digital ATM interface driver provides a command to limit per-VC bandwidth on an interface-wide basis, but preliminary tests indicate that this doesn't seem to work under DU4.0 (revision a); Classical-IP SVC setup signals still request the full ATM network capacity.

Bandwidth between individual EBN sites is determined by the PVP and PVC configuration of the EBN, which is configured manually at Telstra Research Labs. We have no way of affecting, remotely, the configuration of the EBN switches. However, a solution to this problem may exist. Once we have allocated EBN capacity between sites using EBN PVPs, we can then divide this capacity between individual PVCs ourselves. Briefly:

- We are currently using PVCs to communicate over the EBN.
- Our host ATM adaptors can control bandwidth to PVCs. (Tests have shown that this works on our FORE and Digital interface cards.)
- All of our switches will support, and enforce, bandwidth allocation to PVCs.
- We can manage the PVC configuration of our switches (and routers, potentially) using a standard management protocol such as the Simple Network Management Protocol (SNMP) [8].

We could establish a network infrastructure based on PVCs (switched within EBN PVPs), and manage the PVCs using SNMP. We could then explore ways to bind PVCs to specific applications, giving us the means to control the bandwidth used by those applications. This will take some thinking, however, and probably also significant development effort. Three important issues to consider here are:

- We will be implementing a simple proprietary signalling protocol for ATM, where a public standard already exists.
- SNMP is not secure.
- There may be management concerns caused by the DHPC controlling, in an unconventional and experimental manner, switches not owned by ACSys; in particular the University of Adelaide Computer Science department ATM Switch (the ASX-1000).

An alternative is to write our applications using native ATM APIs so that we can directly control the bandwidth used by applications. Unfortunately, there are no widely supported non-proprietary ATM APIs for Unix.

5.4 Connection Limitations

We recently faced the question of how to connect Canberra to more than one other EBN site. Unfortunately, until some of the limitations in our equipment are removed, or SVCs are made available on the EBN, our range of possible solutions is limited.

Without EBN SVCs, we are required to allocate, in advance, how much bandwidth we wish to devote to individual connections between sites - whether these be EBN PVPs or EBN PVCs. This is seen as one of the primary shortcomings in the EBN and Telstra's commercial service, Accelerate ATM. To communicate with two remote sites, each at 10 Mb/s for example, it is necessary to purchase 20 Mb/s of local connection capacity because the circuits to each other site are configured semi-permanently. In this form, ATM is really, from a carrier's point of view, a configurable architecture for providing a leased-line service. This is less than ideal for some networking applications, particularly since switched circuits are already technically feasible which could provide a more flexible service. As users of the EBN, we can't avoid this bandwidth-partitioning problem, so we need to find ways to use the static connections that we do have most effectively.

The following sections describe a few approaches we can take, in order to explain the trade-offs.

5.4.1 Multiple Connections Without Multiple VP Support or SVCs

Until recently, we were using a virtual path through the EBN between the ASX-1000 in Adelaide and the ACSys GIGASwitch in Canberra. This meant that Canberra was restricted to a connection with just one other EBN site because of the Virtual Path limit in Digital's ATM equipment. The Macquarie EBN site recently requested a connection to Canberra through the ACSys EBN Connection, which meant that we needed to consider a new approach. One solution was to arrange a PVC from Macquarie to Adelaide and back to Canberra. This approach had been tried and tested, but it delivered a network connection with very high latency.

We decided that if we relinquished the Permanent Virtual Path between Canberra and Adelaide, then we could use EBN PVCs to establish connections to Adelaide and Macquarie simultaneously, dividing our bandwidth between the Macquarie and Adelaide PVCs. Regardless of how the connection to Macquarie was configured, if we continued to use a mesh of PVCs between Adelaide and Canberra then the EBN would necessarily have allocated a fixed portion of the DHPC bandwidth to each EBN PVC, wasting network capacity. Another down side to a mesh of EBN PVCs was the labour cost. EBN PVCs are established manually at Telstra Research Labs so changes in our configuration would have been inconvenient.

Returning to the our original partially routed configuration using the DECNIS would have improved matters. However, we would still have needed to establish several EBN PVCs to Adelaide, allocating a fraction of the DHPC's share of the EBN connection to each EBN PVC – again, wasting network capacity. We were able to avoid this bandwidth-partitioning problem in our first EBN configuration by using a Virtual Path between Canberra and Adelaide, leaving bandwidth allocation up to the DECNIS router, the ASX-1000 and the Adelaide workstations, which they handled dynamically with flow control.

We decided that each site would provide a router, be it a dedicated piece of equipment or a designated host. Then we could establish a single PVC from Canberra to the routers at each of the two other sites and route all traffic on these two PVCs. We still had to decide how much of the ACSys EBN capacity would be dedicated to each connection, but this problem is inherent in any PVC or PVP service.

This is our current EBN configuration. We consider this to be an interim solution until we have installed the upgrades necessary for our GIGASwitches to support multiple Virtual Paths. The down side to this configuration is the lack of support for non-IP traffic and the load on the machines used as routers. There is also a small, added latency through each router, but this is not significant. The advantages to this configuration are the low maintenance effort, and its reliable performance. We have found that this configuration has so far delivered the best performance between the Adelaide and Canberra sites.

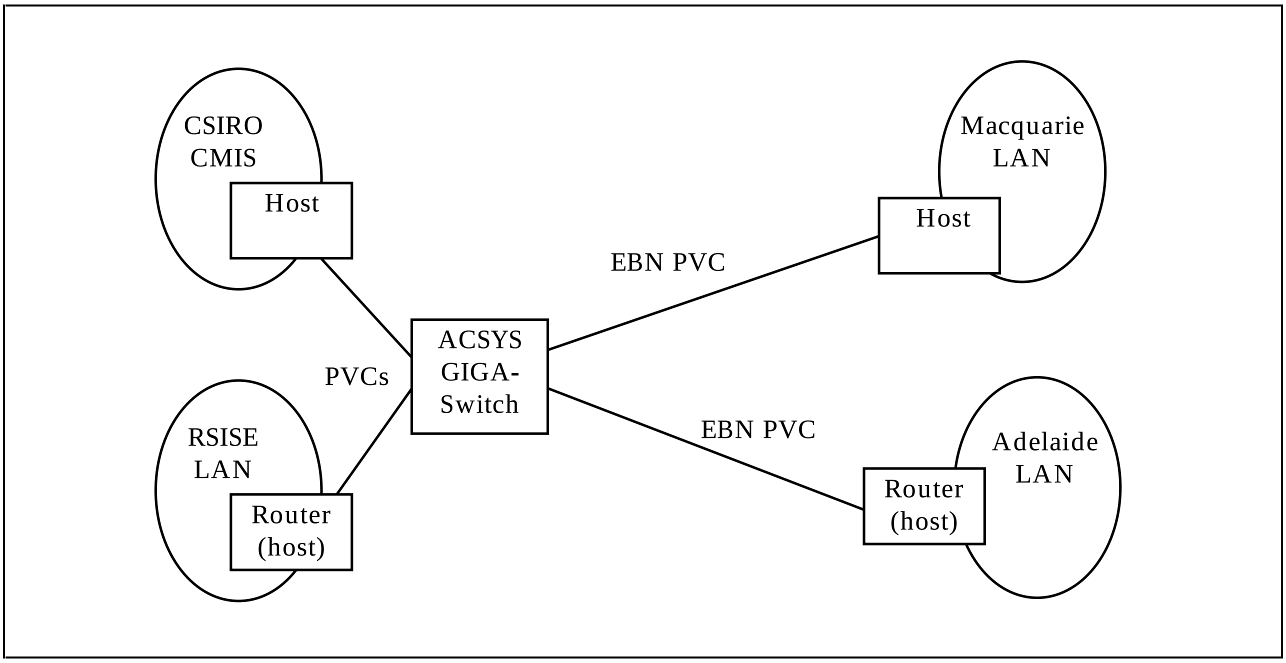

Figure 5-1: Fully Routed Configuration.

5.4.2 Multiple Connections With Multiple VP and SVC Support

With multiple VP support alone, we could do away with routers by making use of meshes of PVCs where required. Because EBN bandwidth can be allocated to a EBN PVP as a whole, we can avoid setting bandwidth on a PVC by PVC basis, leaving traffic management to our ABR flow control mechanism. This solution still requires significant effort to maintain when hosts are added and removed at each site, but this does not involve Telstra Research Labs.

We expect that when Digital delivers multiple VP support, the Macquarie GIGASwitch will correctly support SVCs with the ANU, RSISE and Macquarie GIGASwitches. If this is the case, we can continue to use a PVC mesh between Adelaide and Canberra as long as necessary, while making use of SVCs between the GIGASwitches.

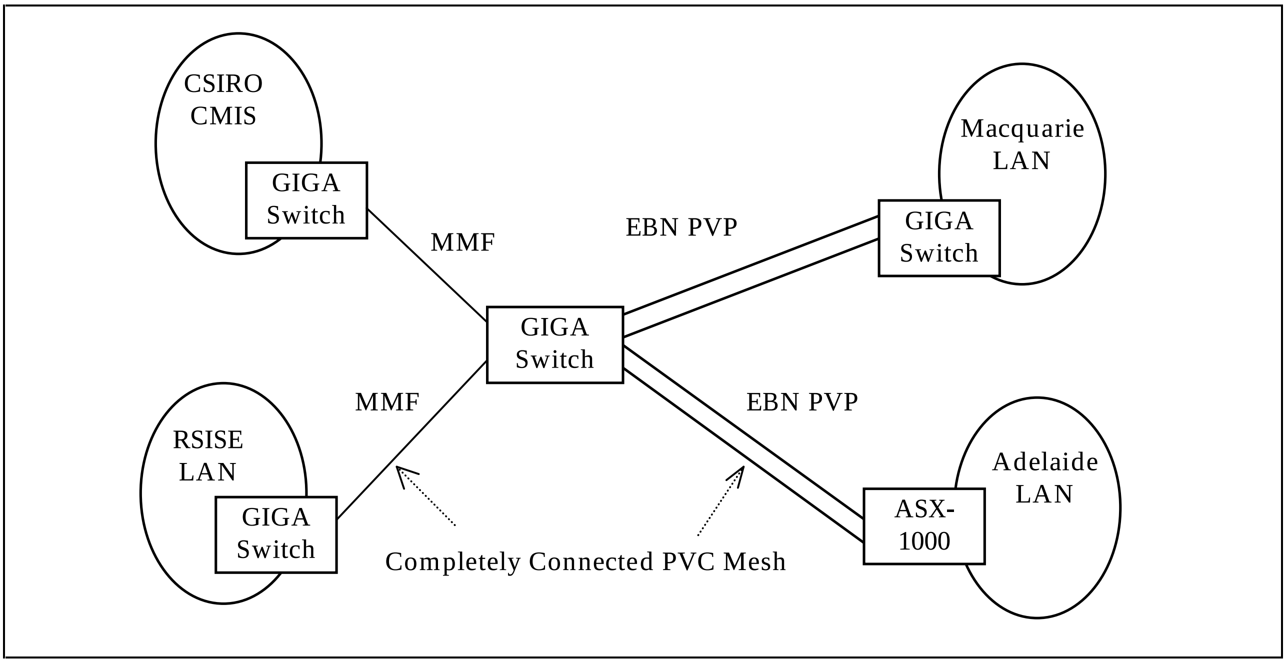

Figure 5-2: Multiple VPs, no SVCs.

The ideal configuration, illustrated in Figure 5-3, will make use of SVCs between Adelaide, ANU and CSIRO in Canberra and Macquarie. This configuration uses no routers and will support native ATM applications if users care to implement these. Before we can implement this configuration, the signalling problem between the ASX-1000 and the GIGASwitch must be solved, and we must implement multiple–VP support and VP-tunnelling for the GIGASwitch.

This configuration should not require significant management effort. Once SVC signalling is in place, configuring Classical-IP to make the system look like one or more private LANs will be no more difficult than setting up our own DHPC Alpha farms (ie, not difficult, and quite reliable). Connections can automatically be established between hosts when they need to communicate, rather than statically configured by a human operator. This configuration might also provide a useful testbed for experiments with LAN Emulation and IP switching.

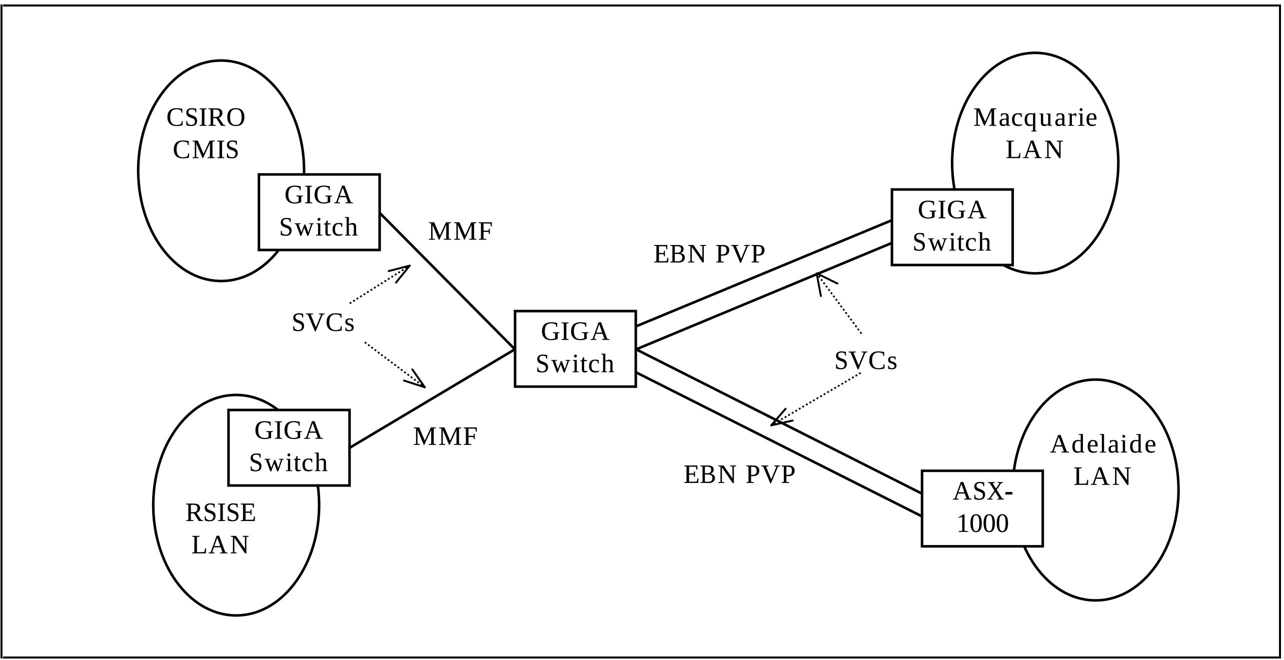

Figure 5-3: Multiple PVPs, SVC Support through Virtual Path Tunnelling.

6. What Can We Currently Achieve With The EBN?

Establishing connections between EBN sites is straightforward, but flexibility is currently limited by the capabilities of our ATM equipment. We have two basic options:

- We can make use of EBN PVCs to connect our sites, and statically allocate our bandwidth between these individual point-to-point connections. This configuration makes good sense if we intend to use routers at the endpoints of EBN PVCs. Routers exclude non-IP (raw ATM) traffic, but we could can establish extra non-IP PVCs when they are required to support raw ATM applications.

- We can use EBN PVPs to connect our sites and achieve more flexible connections by establishing our own PVCs and using ABR connections and flow control to manage bandwidth. When we have overcome some problems with our equipment, we also hope to implement Switched Virtual Circuits on top of EBN PVPs.

In either case, we will have to partition our total EBN connection bandwidth between individual site-to-site connections, because the EBN does not provide bandwidth-on-demand in the form of SVCs.

Network management (referring to connections between sites, and dynamic and static bandwidth allocation) is currently quite inflexible. We have not developed any mechanism to manage network resources at the application level, and no standard framework exists that we are aware of. Our network configuration is currently semi-static. Operator intervention is required to change any aspect of the configuration. In section

5.3, we have suggested one possible approach to implement better network management using the facilities we have available. This solution is fairly clumsy, and needs more thought.

In spite of the limitations, we have obtained good IP network performance over the EBN using three configurations – two using routers, and another using a mesh of Permanent Virtual Circuits. We have found that flow control is a significant issue in ATM networks, and that the common flow control standards, EFCI and TCP slow-start, are not particularly effective at managing traffic across wide area networks such as the EBN. We have also investigated ways to try to minimise the effects of traffic burstiness.

7. Desirable Qualities In ATM Networks

There are a number of features that would be desirable in any future wide-area ATM service.

- The first is access to higher bandwidth than 34 Megabits / second. We would also prefer to utilise all of the bandwidth that our connection offers us. The 80% restriction on our E3 link means we lose around 6.8 Mb / second of EBN throughput. This figure gets worse as the capacity of our link increases.

- The second feature we would like to see is support for SVCs, as well as the current PVP and PVC services. SVCs give us access to bandwidth on demand. SVCs also provide a way to make incidental connections between sites with no prior permanent connection, supporting short–lived applications such as teleconferencing and voice connections. PVP and PVC services remain useful, as they support the use of equipment with no SVC capability (particularly some ATM-equipped routers) and allow dedicated VP connections for implementing private SVC-based networks (this would be extremely useful when a customer wants to employ a proprietary switching protocol such as FORE's SPANS protocol or Digital's FLOWMaster).

- Another feature we would like to see in a wide-area ATM network is a degree of tolerance for bursty traffic – the more the better. Unfortunately, with our limited experience it's difficult to quantify this requirement. More generally, support for a range of traffic types (CBR / VBR / ABR) is desirable.

- Support for multicast PVCs will benefit video conferencing applications. We can implement multicast switching in our own equipment (certainly the GIGASwitch, and possibly also the ASX-1000) but this requires us to allocate, from the sender's WAN connection capacity, the data rate for each outgoing branch of the PVC. If the carrier network provides multicast support, then the sender need only allocate the bandwidth for the trunk of the PVC - leaving the branches to the carrier.

It is perhaps also worth mentioning a few desirable features in customer-premises ATM equipment.

- Support for SVC signalling.

- Support for multiple Virtual Paths.

- A sophisticated cell buffer implementation. Digital and FORE provide good cell buffers, Cisco (at the time of writing) don't support per-VC buffering which means that one connection can cause cell-loss in another Virtual Circuit.

- Support for Multicast Virtual Circuits.

- Ability to snoop cells in transit.

A less important issue is switch transit speed – commercial ATM switches can switch cells on the order of a few microseconds. FORE currently has one of the fastest implementations, but Digital's GIGASwitch / ATM is only a few microseconds behind [21]. What will be more important, however, as networks scale, are the port buffer implementations and the performance of the switch control processor, which affects the rate at which SVC can be established and the provision of services such as LAN Emulation. FORE and Digital both perform well in this area.

8. Acknowledgments

Most of this work was undertaken with the supervision and assistance of Dr. Mike Rezny, who has since left the project. Markus Buchhorn has also been of enormous assistance in the design, configuration and management of our EBN connections and ATM equipment.

The authors wish to Acknowledge that this work was carried out within the Cooperative Research Centre for Research Data Networks, established under the Australian Government's Cooperative Research Centres program.

The Distributed High Performance Computing Project (DHPC) is a project of the Research Data Networks Cooperative Research Centre (RDN CRC), is managed by the Advanced Computation Systems CRC and is a joint activity of the Australian National University and the University of Adelaide. We thank Telstra for the provision of the Experimental Broadband Network, and we particularly would like to thank Daniel Kirkham at Telstra Research Laboratories for his invaluable help.

9. References

[1] D. Kirkham, "Telstra's Experimental Broadband Network", Telecommunications Journal of Australia, Vol 45, No 2, 1995.

[2] W. Richard Stevens, "TCP/IP Illustrated", Prentice-Hall 1994, ISBN 0-20-163346-9

[3] W. Richard Stevens, "Unix Network Programming", Prentice-Hall 1990, ISBN 0-13-949876-1

[4] M. Laubach, "Network Working Group Request for Comments: 1577, Classical IP and ARP over ATM", 1994.

[5] J Heinanen, "Network Working Group Request for Comments: 1483, Multiprotocol Encapsulation over ATM Adaptation Layer 5", 1993.

[6] "LAN Emulation Over ATM Version 1.0", The ATM Forum Specification af-lane-0021.000, 1995.

[7] J. A. Matthew, K. A. Hawick, "Applying ATM to Distributed and High-Performance Computing on Local and Wide Area Networks", Technical Report DHPC-016, The RDN CRC Distributed High-Performance Computing Project.

[8] J. Case, M. Fedor, M. Schoffstall, J. Davin, "Network Working Group Request for Comments: 1157, A Simple Network Management Protocol (SNMP)", 1990.

[9] W. Stevens, "Network Working Group Request for Comments 2001: TCP Slow Start, Congestion Avoidance, Fast Retransmit, and Fast Recovery Algorithms", 1997.

[10] "Frequently Asked Questions on ATM and Digital's ATM Program", Digital Equipment Corporation 1996. Available online via http://www.networks.digital.com/dr/techart/

[11] "ATM User-Network Interface Specification version 3.0", The ATM Forum Specification af-uni-0010.001, also pub. PTR Prentice-Hall 1993, ISBN 0-13-225863-3

[12] "ATM User-Network Interface Specification version 3.1", The ATM Forum Specification af-uni-0010.002, approved 1994.

[13] R. Handel, M. N. Huber, S. Schroder, "ATM Networks - Concepts, Protocols, Applications." Addison-Wesley 1994, 0-20-142274-3

[14] M. Batubara and A. J. McGregor, "An Introduction to B-ISDN and ATM", Technical Report 93/4 Faculty of Computing and Information Technology /Department of Robotics and Digital Technology.

[15] Paul Reilly, MSD Marketing Silicon Graphics Inc. "PDH, Broadband ISDN, ATM and All That: A Guide to Modern WAN Networking and How it Evolved", Silicon Graphics 1994. Available at: ftp://sgigate.sgi.com/pub/Surf/ATM.ps.Z

[16] Cisco hosts an excellent online reference for a range of networking and internetworking technologies, including ATM. http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/index.htm

[17] M. D. Wilson, S. R. Taylor, M. Rezny, M. Buchhorn, A. L. Wendelborn, "ACSys/RDN Experiences with Telstra's Experimental Broadband Network - First Progress Report", Technical Note DHPC-022, Advanced Computational Systems CRC, 1997

[18] T. Des Jardins, S. S. Sathaye, "Traffic Management in FORE Systems' ATM Networks", (DRAFT) Doc#8.4.2, Fore Systems Inc, 1994.

[19] "Traffic Management Specification Version 4.0", The ATM Forum Technical Committee specification af-tm-0056.000, April 1996

[20] R. Mandeville, "The ATM Stress Test: Which Switches Survived?", Data Communications Magazine, 1995.