CONTENTS

SPECIAL ISSUE ON COMPUTER NETWORKS

INDUSTRY CASE STUDIES
34-41 Transport Australia Packet Switching Network
   W.T. MALONEY
42-45 Introduction of a State Wide Computerised Network
   A.J. WINDROSS

NETWORKS FOR EDUCATION AND RESEARCH
46-51 Network Developments at the W.A. Regional Computing
   Centre, 1965-1980
   T.A. REID
52-57 The Sydney UNIX Network
   R.J. KUMMERFELD and P.R. LAUDER

RESEARCH PAPER
58-63 Decentralised Scheduling
   L.M. CASEY

SHORT COMMUNICATION
64-65 HYPERDISK, An Access Method for Remote Disk Devices
   G. HAVAS

SPECIAL FEATURES
33 Guest Editorial
66-67 Letters to the Editor
68 Future special issues
"This wouldn't have happened, Douglas, with a Tandem NonStop™ System."

**When you lose your computer, do you lose your customers too?**

Your customers count on you to deliver. If you don't, they may disappear. Sometimes forever.

All it takes is one small failure somewhere in the system. One disc or disc controller. One input/output channel. Something you've never heard of can put you out of business. Unless you have a Tandem NonStop™ System.

**NonStop™ Operation, to be sure.**
This system keeps on running, right through a failure which would shut down any other system in the world. Including a failure in a disc, a disc controller, an input or output channel, even one of the processors.
No one else can approach this performance.
Tandem takes even a critical failure in stride, without loss or duplication of any transaction-in-process.
No downtime. No lost records. No lost customers.

**The company's worth keeping.**
Especially your own. And you'll be joining excellent company, too.

The Tandem Users Group reads like a "Who's who" in international business, industry, finance and government. You don't have to be a member of the Fortune 1000, but most of the companies with Tandem NonStop™ Systems installed are in fact the leaders in their fields. They looked and they bought Tandem. The system, the software, the support. And the company behind it.

**The way to grow.**
Without penalty and without regrets. Tandem NonStop™ Systems let you start with only the computer system you need right now. As you need them, you can add processors, terminals, programs and locations to take you from a local system to a global network. And never lose one cent on your original investment. Same hardware. Same software. With low cost additions, not replacements. And you'll be keeping your customers, too.
"News Briefs from the Computer World" is a regular feature which covers local and overseas developments in the computer industry including new products, interesting techniques, newsworthy projects and other topical events of interest.

ELECTRONIC AIR CLEANERS FOR OFFICES

The conference room meetings held in a cloud of cigarette smoke need no longer be the case say the manufacturers of a new range of electronic air cleaners. Honeywell are marketing a range of electronic air cleaners specifically designed for the removal of cigarette smoke from offices.

Depending on the size of the room Honeywell say they have a unit to match the need. Their new self-contained F54D model comes complete with a 3 speed fan which enables the unit’s capacity to be selected for the room size or smokers present. The unit can be wall or floor mounted or fitted with castors for moving from room to room.

It plugs into a standard power point enabling its electrostatic precipitation principle to filter up to 95 percent of cigarette smoke, dusts and pollens that pass through it.

Units suitable for servicing areas of 30 square metres up to 130 square metres in size are available. For larger areas combinations of self-contained units or ducted systems can be arranged.

JONES & RICKARD 60 HZ SHORE SUPPLIES FOR RAN

Jones & Rickard Pty Ltd of Sydney are supplying seven 250 KVA, 50/60 HZ Frequency Changers to specifications issued by Department of Housing and Construction. Two sets are to be installed at HMAS Moreton in Brisbane, two at HMAS Waterhen in Sydney and three at the new RAN Patrol Boat Base in Cairns. The sets are to provide 60 HZ shore supplies to patrol craft and other RAN vessels, to avoid the need to run the ship’s engines to generate this type of power when the vessels are moored at the wharves.

Each of the frequency changers comprises a 415 volt, 3 phase, 50 HZ, 600 RPM synchronous motor on a common shaft with a 250 KVA 440 volt, 3 phase, 60 HZ brushless alternator. The frequency changers are designed and manufactured at Jones & Rickard Sydney works. The sets are required to operate in parallel, and the control cubicles incorporate equipment for automatic synchronising. The equipment is of high electrical performance necessary to comply with the Australian Defence Standards.

DIGITAL EQUIPMENT OPENS CHATSWOOD OFFICES

Digital Equipment Australia has consolidated the majority of its Sydney operations under the one roof in the Northern Tower of Chatswood Plaza, Chatswood.

In a ceremony attended by invited customers, press and Digital Equipment staff, the offices which comprise all twelve floors of the Tower Building were officially declared open by New South Wales Premier, Mr Neville Wran.

Previously the computer company’s Sydney operations and South Pacific headquarters were dispersed through seven locations. Now the customer can ring through to sales and services on the one number.

50 years covers a lot of experience in ELECTRICAL REPAIR SERVICES

FROM SMALL MOTOR REWINDS TO THE LARGEST HIGH VOLTAGE ROTATING ELECTRICAL MACHINERY

NEW $1 MILLION FACTORY EXTENSION APPROXIMATELY DOUBLES THE EFFECTIVE WORK AREA UNDER THE 15 AND 5 TON CRANES

- Fast small motor repairs (24 hour service when needed)
- Specialised equipment for repair of lifting magnets
- Authorized repair shop for Defence Forces and electricity Supply Authorities
- 6.25 MVA, 11KV alternator stator repair

Our facilities include:
- 15 tonne lifting capacity
- Dynamic balancing up to 15 tonne weight
- 18m vertical boffer and comprehensive machine shop
- Shaft grinding and repairs using MIG shaft build-up
- 18m vertical boffer and comprehensive machine shop

GENERAL ELECTRIC Franchised Service Shop

JONES & RICKARD PTY. LTD.
889 South Dowling Street, Waterloo, Sydney, N.S.W. 2017. Telephone: 663 3938 Telex: AA25668 JR84A

SCHOOL LEAVERS TRAIN AS COMPUTER SERVICE ENGINEERS

A pilot scheme to train school leavers in the maintenance and servicing of computer equipment has been introduced by IBM Australia.

In the initial intake, six school leavers, five boys and one girl, have begun 12 months training as Student Customer Engineers. Two are located in Melbourne and four in Sydney.

"Response to the recruiting in Sydney and Melbourne was excellent," said Graham Spong, Director of Data Processing Customer Engineering, "indicating a high level of interest in this type of career among school leavers".

Mr. Spong said the decision to implement the student training programme was based on the changing needs of the service environment and the fact that some school leavers, given the right amount of training could meet some of those needs. In addition, the scheme provided an opportunity for young people to get started in a career which otherwise required tertiary qualifications and/or experience.

In selecting the first batch of trainees, the company had looked first for students with a science/physics orientation. In the final selection, presentation was a significant factor to identify the sort of people who looked to be self starters.

During their 12 months training, the recruits will gain experience in all aspects of the customer engineering operation — including the Maintenance Parts Centre, at Waterloo, NSW, the Customer Support Centre, at Rosebery, NSW, and working with experienced customer engineers in the field.

As they gain experience they will assist with installation of equipment. Another facet of their training will cover customer relations — learning to develop "people skills".

FULL COLOUR ID CARDS IN 29 SECONDS

The Dek/Electro System XII photographic identification system produces inexpensive, full-colour ID cards to aid company or institutional security.

System XII, distributed in Australia by P.D. Security Pty Ltd, is a compact unit containing camera, lighting strobe and die cutter. Identification cards can be produced quickly by one operator in a few simple steps: the operator is led through the photographing procedure by a lighted sequencing logic which is electronically interlocked to prevent operator errors. Easy daylight loading of film packs is a standard feature of the unit.

System XII produces a bright, full-colour image in only 29 seconds — the card emerges from the machine protected by a plastic laminate and, unlike other machines, heat encapsulation or additional lamination is not required for further security. The ID cards may be used to provide immediate, non-transferable identification. Standard security features, such as validation signatures and seals, are included in the system.

The Dek/Electro System XII can be integrated with all existing photo identification systems, and will accommodate the wide range of form sizes and formats produced by data collection and printout systems. The system may also be supplied with a built-in negative system, to allow film records to be retained on negative.

System XII is part of a wide range of Dek/Electro identification systems which has proven highly successful in North America — more than seventy percent of all colour photo driver licences in the US are produced with Dek/Electro equipment.

Complete details on the Dek/Electro System XII are available from the Australian distributors P.D. Security Pty Ltd, Artarmon, NSW.

AUDITORS COME TO GRIPS WITH COMPUTERS

The Institute of Chartered Accountants in Australia has developed a comprehensive series of computer audit courses for the 1980's.

The courses cover the requirements of both public accountants and internal auditors, from those who are inexperienced in computer auditing to those with considerable experience in the field.

Auditors will be offered a planned programme of development which will enable them to audit computer systems of varying levels of complexity.

The new series will be launched with the introduction of three courses — computer controls, computer auditing and auditing advanced computer systems. It is anticipated that the courses will be run in capital cities this year.

These courses are the product of a committee of computer audit experts and each course will be presented by computer audit specialists from well-known accounting firms.

For further information please contact Marg Donnan, Professional Development Office, telephone (02) 290-1344. (Continued on Page III at the back)
This issue of the *Australian Computer Journal* is devoted to the topic of computer networks, with particular emphasis on developments within Australia. The selection of papers presented here does not tell the whole story of what is happening locally, but it does allow some deductions to be made. Computer networks are being developed and used, but they are almost all for specialised purposes and, for the most part, limited by the service offerings of Telecom Australia, which has the monopoly on terrestrial telecommunications. Since the unhappy episode with its Common User Data Network, Telecom has been reluctant to pioneer in new developments in data communications. Australian computer users today find themselves in a large country, with widely spaced centres of population, paying high costs for leased intercity data communication channels, and still without a public data communications network. At present the nearest approach to such a network is CSIRONET whose design was originally predicted on the assumption that one party to every conversation would be the CDC 7600 computer located in Canberra, with multi-host supported developed later.

It will be most interesting to follow events from 1982 when Telecom is due to begin its public packet-switched service. Certainly, if the recent experiences of the Overseas Telecommunications Commission with its MIDAS service (see the *Australian Computer Bulletin*, November 1980) can be used as a guide, public demand for this service will develop very rapidly. Computer networking may then come of age on the Australian scene.

Mr Maloney’s paper describes a purpose-built package-switching network for messages associated with Australia’s air navigation system. Built as a special purpose network because there was no reasonable alternative, it forms part of a larger international network conforming to CCITT standards.

In North America and Europe, pioneering computer networks were developed for airline reservation and integrated banking systems, but in Australia, it has been automated off-course betting systems that have led the way. Mr Windross’s paper describes some of the experiences of the NSW Totalizator Agency Board in extending its services throughout the State.

Mr Reid’s paper provides an account of the history of one of the early regional networks developed to provide distributed access to one or more centrally located mainframe computers. As the emphasis on access to raw computer power lessens, the network is evolving towards a new role in information transfer, with connections to many other machines and other existing networks.

Dr Kummerfeld’s and Mr Lauder’s paper describes an interesting local development of a network which connects more than 20 minicomputers located on two separate university campuses. The network design frankly exploits the capabilities of the operating system used by the interconnected machines and, developed on a minimal budget, its effectiveness and unconventionality challenge some of the current wisdom as to how public networks ought to be constructed.

Dr Casey’s paper is not tied to a particular system. It explores some of the problems that arise in distributing computational tasks among different processing elements of a load-sharing network. Details are given for a proposed load balancing scheme that minimises bandwidth requirements.

Finally, in a short note, Dr Havas describes a further development of the CSIRONET Local Computer Network, namely an access method for disk devices.

**BIOGRAPHICAL NOTE**

John Lions is currently Associate Professor of Computer Science at the University of New South Wales. He was born in Sydney in 1937 and received BSc and Ph.D degrees from the University of Sydney and the University of Cambridge. From 1963 to 1970, he worked in Canada, first as a management consultant in Toronto, and later as Director of the Computer Centre at Dalhousie University. After spending two years with Burroughs Corporation in California, he joined the University of New South Wales in 1972. His interests include operating systems, computer networks and timetabling problems.
The Commonwealth Department of Transport is currently involved in the development of a modern packet switching network for interchange of aeronautical data. The project has involved considerable hardware and software design, in addition to the usual station design. An overview of the design philosophy is presented and the hardware and software approach to implement the network is described. The importance of international standards and their application in the national environment is discussed.

Keywords and phrases: Communication standards, computer network, packet switching.
CR Categories: 3.29, 3.57, 3.81

1. INTRODUCTION
Transport Australia operates an extensive message switching network as part of the world-wide Aeronautical Fixed Telecommunication Network (AFTN). The system carries operational data associated with aircraft movements, and some 10 million messages are currently delivered annually to domestic and international addresses. Ten categories of message may be transmitted over the AFTN including:

- Flight movement and control messages (flight plans, departure advice, etc.).
- Notams (Notices to Airmen).
- Meteorological messages (terminal forecasts, route forecasts, etc.).

The existing network has been largely based on low speed telegraph techniques with manual procedures for supervision. To maintain a reliable service and minimise trunk requirements the Australian AFTN has involved a number of switching centres located in major areas. Over the past 10 years an extensive automation program has been implemented to replace manual tape relay centres with computer or firmware based switching systems, and now 70 per cent of all channels are connected to automated centres. There are 400 connections serving some 50 Departmentally staffed stations and a similar number staffed by other agencies (RAAF, Met, etc.). Distribution is extensive at locations such as major airports, where over 30 terminals may be required.

2. USER FACILITIES
Although automation of AFTN facilities world-wide has significantly improved the operation of the AFTN, it has been clear in recent years that a major change in network performance would be required to meet future requirements for data interchange by the aeronautical community. The integrity and response time of the low speed network has been found inadequate to support widespread application of data processing in Air Traffic Service applications. Various studies of the implementation of air-ground data links have indicated that, other than in specialised applications, widespread development depends very much on the availability of a high integrity network to access the required data.

It was also recognised that the fundamental role of the AFTN as a 'store and forward' message switching network would remain important, in addition to those new facilities and capabilities a modern data network may offer. In considering the development of data networks a broad division may be made between data interchange 'transport' functions, and application functions which use the data network as a carrier. Some of the important communication type user facilities at present available include:

- Group Address:
  - a single address identifying a number of addresses.
- Multiple Address:
  - multiple individual addresses.
- Message Priority.

Where the communication function is provided as a front end to a host computer, a number of applications of importance to civil aviation may be considered, such as:

- ATC Computers:
  - data exchanges between processors.
- Flight Information Systems:
  - local and remote access to aeronautical files.
- Flight Briefing Service:
  - various types of processing accessing FIS system files.
- Flight Plan Systems:
  - storage and calculations associated with flight plans.
- SAR Systems:
  - calculations associated with Search and Rescue.
- Technical Management Systems:
  - centralised fault statistics, fault reporting, etc.

3. STANDARDISATION
The International Civil Aviation Organisation (ICAO) is responsible for the development and uniform application by Contracting States of international standards pertaining to civil aviation. National regulations or practices applied within a State may, of course, differ from the international case, but there are strong reasons for main-
Packet Switching Network

The body entrusted by ICAO to develop technical standards for international aeronautical data interchange is the Automated Data Interchange Systems Panel (ADISP) in which Australia has maintained an active participation. Over recent years the ADISP has been developing data communication transport standards for a Common ICAO Data Interchange Network (CIDIN) to serve the aeronautical community. These standards are based on packet switching techniques, which are a logical extension of the 'store and forward' principle on which the existing AFTN is based. Advantages of packet switching in a private network may be summarised as:

- Economic use of network by dynamic resource allocation.
- High availability with efficient alternate routing and minimal transit delay.
- Ability to handle multiple users and different codes and formats.

The ADISP has been progressing its task at the same time as the standards work currently underway in bodies such as the ISO and CCITT. This has enabled close alignment of the ADISP standards with the approaches of other international organisations, with ICAO specific requirements and options being defined as necessary. The Panel has been concentrating on the definition of 'data transport service' type functions and has adopted the layered structure now emerging as the reference architecture for data systems. Of the seven levels in this structure the first four have been considered. These are:

- Level 1 - Physical Interconnection.
- Level 2 - HDLC Link Protocol.
- Level 3 - Network Control Procedures.
- Level 4 - Transport Procedures.

In Level 4, ICAO requirements for facilities such as Multiple Dissemination, User Acknowledge and Priority are specified. These functions are provided over and above any capability existing in the Network layer (Level 3). It should be noted that CCITT standards X.25 and X.75 relate to the interface specification between user equipment and packet switching networks or between packet switching networks. The work of the Panel is concerned with formats and procedures to be used within a packet switching network, as well as the capability to interface to other private networks and public data networks.

4. DOMESTIC DATA INTERCHANGE PLAN

Transport Australia has been planning new AFTN facilities for some time, with the objective of satisfying a number of pressing technical and operational requirements. Consequently a five-year plan has been prepared to provide data type facilities at all stations. An important aspect of the plan is the method of evolving existing domestic procedures to achieve essential compatibility with the new ICAO standards. Fundamental hardware and software constraints limit the extent to which existing computer installations may be upgraded. In essence, the plan provides for packet switching facilities at 37 locations and data facilities using upgraded character link procedures at 19 locations.

The first application of the packet switching network involves a major upgrading of the Western Australian AFTN, using a system known as Dataflash, developed within the Planning, Research and Development Branch of the Department. Installation of this system is scheduled for the latter half of 1981 and follows delivery of equipment and associated development over the past two years. The Western Australian project involves installation of equipment at 12 locations including a network control centre at Perth airport. The network will support 55 terminals in various operational areas. A diagram of the proposed packet network is shown in Figure 1. Remaining low speed channels are not shown. As with other areas of the network, stations are interconnected by either Departmental or leased bearers. The aim is to provide a high degree of connectivity using physically alternate paths, and this feature will be enhanced as the network is developed over the plan period.

5. NETWORK TOPOLOGY

In determining the design philosophy and implementation approach for the network a number of inter-
related factors were considered, the principal being:

- Development at a time of rapidly changing technology.
- High reliability and availability requirement.
- Provision and supervision of new facilities.
- Interface to existing network procedures.
- Applicability of future public packet switching service.
- Cost effectiveness.

It was concluded that an approach that combined commercially available hardware with specially developed modules, along with in-house software development of the new procedures, was most appropriate. Due to the time lag between the emergence of standards and the manufacturer's response in providing suitable equipment, special interfacing arrangements are to be expected.

The comprehensive facilities and reliability required throughout the network suggested a distributed processing approach, while the requirement to interface vastly different procedures was ideally suited to a supervisory centre based on a multi-processor organisation. The Network Control Centre (NCC) design adopted is based on tightly coupled multi-processors for the communication function, with provision for a loosely coupled data base facility. This offers a flexible and powerful processing capability to accommodate virtually any application task likely to arise on the domestic network. The multi-processor approach enables functionally similar tasks to be handled in individual processors, with a central processor handling common routines and AFTN application tasks. This is attractive when integrating new procedures into an existing network.

An integral part of the concept is the provision of Automatic Terminal Controllers at each node of the network. These provide the function of remote data terminal interface and packet exchange. Terminal controllers have considerable autonomous processing capability and satisfy the design aim of providing similar facilities at remote stations to those provided at the centre. Terminal controllers are accredited to a designated supervisory centre. As the network is expanded additional terminal controllers may be accredited to another NCC. When selecting equipment on which the network would be based, particular attention was paid to ensuring a high degree of software and hardware compatibility.

6. NETWORK CONTROL CENTRE HARDWARE

A block diagram of the Network Control Centre is shown at Figure 2. The system is based on the Texas Instruments' 990/10 mini-computer, and features in the architecture of this computer are an essential facet of the implementation. The 990 series provide dual input/output systems. The Communications Register Unit (CRU) is a serial input/output port which offers a simplified method of inputting/outputting data at moderate speed for communications or control applications. The Ti-line is the high speed asynchronous DMA bus and has a transfer capability of 50 Mb/s.

The Ti-line offers an ideal method of interconnecting multi-processing systems. For the data interchange function three processors are provided, each processor handling communication links requiring similar formats and procedures. These line processors are each equipped with 32K words of error-correcting memory. The central
processor communicates with the line processors over the Ti-line, to a standard transfer procedure as described later. The central processor contains 96K memory and the operating system extends over all processors. This system can therefore be considered tightly coupled. A 25 M byte disc drive is connected to the central processor, which controls disc handling. The database is connected over a further Ti-line coupler to the central processor; however, this may be considered solely as a communication link. All access requests to the database will be via the data transport system. The TTY interface is provided for diagnostic work on each processor when the system is off-line and comprehensive facilities are provided for technical supervision.

On the communication link side of the system, it is necessary to provide an efficient means of interfacing large numbers of channels. This has been achieved by designing interface modules that support up to eight duplex channels for both HDLC links and data circuits. These then interface through the CRU to the line processor. Line terminating equipment, not shown on Figure 2, provides the external line connection and this equipment contains the input paralleling and output changeover facility for the duplicated system.

The main Dataflash equipment occupies five cabinets. Two cabinets are required for each processing suite (main and standby), with a further cabinet for line interfacing and test equipment. Some of the in-house designed computer modules are housed in the 990/10 processors themselves, while additional modules are housed in separate chassis in the line interfacing cabinet. Separate provision is made for housing modems used on leased trunks, and for a maintenance facility which contains basic interfacing, line processing and central processing test bed. Disc drives are console mounted.

7. AUTOMATIC TERMINAL CONTROLLER HARDWARE

The terminal controller is based on the Texas Instruments’ 990/4 micro-computer. The 990/4 uses the Tl 9900 microprocessor which operates at a clock speed of 3 MHz. It is upward compatible with the 990/10 machine which has additional features such as memory mapping and extra interrupts. The unit is housed in standard six slot chassis with 12K word parity memory and power supply. Terminal controllers use the same in-house designed interface and control modules as used for the centre. Provision has been made for duplication of the facility, using a manual changeover. Dataset and changeover modules are housed in a separate chassis.

8. SPECIAL MODULES

A number of special modules, not commercially available, have been designed to enable the Dataflash concept to be implemented. The modules either mount in the processor chassis or in a separate interface chassis. Modules generally accommodate eight channels on a standard size card. Details are as summarised below and the UART module is briefly described:

- Eight Line CRU Interface (asynchronous):
  - Provides eight data interfaces to data/telegraph circuits and up to 12 digital I/O functions for console presentation.
Packet Switching Network

- Eight Line CRU Interface (synchronous):
  - Provides eight data link control interfaces for HDLC trunk circuits operating at up to 9600 bps.
- Line Telegraph Adaptor:
  - Provides eight Rx/Tx interfaces for telegraph lines operation to 300 bps.
- Dataset Module:
  - Provides standard interface control signals for eight channels in the local airport environment.
- Line Switching Relay:
  - Switches eight output lines on Dataflash/ATC for changeover.
- Control Module:
  - Provides watch-dog timers to monitor operation of the dual installation and miscellaneous functions.

The asynchronous data module is based on a 9900 microprocessor. The card is interrupt driven and contains a RAM buffer through which a 16 bit transfer word is passed to or from the line processor. The card can support eight channels and 12 function controls. Each channel interfaces via a UART chip (TMS 9902) to the V.24 signals. Incoming data is assembled into a character within the chip, and an interrupt signal generated to the on-card 9900 microprocessor. After any higher priority interrupts have been processed the 9900 reads the character from the appropriate chip and stores it in RAM. This eight bit character will then be assembled into the 16 bit Transfer Word referenced earlier with additional address information and loaded in the Output Buffer RAM. The card then generates an interrupt to the line processor which transfers the word with a 'Store CRU' instruction. Alarm conditions are handled in a similar process using the transfer word. The reverse process occurs for data and control output signals. At system initialisation each UART may be set for its correct data rate, character length and parity as determined by program held in PROM. The operation of the HDLC interface module follows that of the asynchronous module. The principal difference is the use of the synchronous TMS 9903 chip which provides flag generation, '0' bit insertion and frame check sequence functions. This chip requires two additional interface lines (receive and transmit clock). The maximum data rate for the interface cards is determined by the on-card software, of which there are several different versions for the different types of lines (AFTN, Telex, etc.). The program occupies approximately 1K byte (PROM) and tests indicate the card can support up to eight channels operating at 9600 bps.

9. DATA INTERCHANGE PROCEDURES

The Dataflash network procedures follow the standards work as closely as possible. The physical interface uses the industry-accepted V.24 standard, equipment configured for the more recent digital interface standard X.21 not being readily available. At the link level the balanced link access protocol (LAP B) is being used. This is a particular case of the generalised High Level Data Link Control (HDLC) procedure specified by ICAO for use on aeronautical data links, and enables link level compatibility for future X.25 interfacing to the public data network.

Network connectivity is based on use of Permanent Virtual Circuits (PVCs). A virtual circuit is an association of so-called logical channels on physical circuits in a network for a period of time. In the initial implementation these will be configured between all nodes in the network and the Perth Centre. To maintain an adaptive and flexible routing structure, tables are configured to enable up to three pre-determined alternate paths (PVCs) between any two locations. Only one PVC may be active at one time for the transfer of operational traffic. Network formats and procedures are similar to those of the CCITT interface standard X.25, but operate in a symmetric mode compared to the asymmetric nature of the interface between a DTE and DCE.

A PVC may be either active, available or unavailable, and may only be activated by the NCC. However, any node may cause the PVC to become unavailable (e.g., due to link failure) or restore PVC availability (following link restoration). The state to be entered is determined by the diagnostic code included in a 'Reset Request' packet. A simple flow control procedure may also be applied on the network. In addition to the non-selective mechanism existing at the link level, selective flow control on a per PVC basis is applied. This limits the number of outstanding packets between adjacent nodes to the window size and applies end to end when delivery confirmation is required. The window is based on the packet send and receive numbers contained in the packet header.

In the case of the Australian AFTN it has been decided not to implement Transport Layer (Level 4) functions at this stage. This implies various user specific...
functions are handled as application tasks by processing in the NCC. Delivery confirmation is being implemented using the 'D-bit' capability of Level 3. The centralised supervision and control functions in the NCC have a two-fold purpose — supervision of the several different procedures interfaced at Perth, and handling of application tasks associated with message flow over the AFTN. The NCC presents the status of all terminals/links connected to the network. The more important functions of the centre are:

- Gateway packet exchange.
- Interface to low speed AFTN network/auto Telex.
- Interface to character procedures (existing net).
- Database access.
- Network status/diversion control.
- Pre-format entry/modification.
- Group address/multi-address process.
- Message retrieval (AFTN regulatory requirement).

In the initial implementation the use of PVCs limits interactive operation to transactions between terminals on the network and Perth. This satisfies the primary objective of message preparation assistance to remote traffic entry, and pre-format access. The next stage of the project will introduce Switched Virtual Circuits (SVCs), providing conversational mode operation between terminals and the ability to establish X.25 connections. The initial use of PVCs simplifies processing as the 'Call Establishment' and 'Clear' phases are not required.

Figure 5. Dataflash software structure

An important application of the network is the access to aeronautical information, and provision has been made for an integrated database in the system configuration. Aeronautical databases consist of a number of subsystem files (e.g., weather, NOTAM) accessible by various keys (location, area, flight route, etc.). The provision of these facilities replaces existing manual methods, enabling efficient data collection and dissemination of aeronautical information. Standards have not yet been set by ICAO for database access formats and procedures, which are identified with the higher layers of the reference model architecture. Work is proceeding on the Australian implementation, although this facility will not be available when the network is commissioned.

10. CONTROL CENTRE SOFTWARE

In the Dataflash centre there is a fundamental difference between the message handling procedures and the procedures for handling transit packets associated with interactive transactions. Handling of message traffic is more complex, requiring additional processing and entry into applications software. Figure 5 shows the main program modules of the system. In the Dataflash system non-interactive message traffic is required to be available for retrieval at the first switching centre (Perth) in the network. The concept adopted has been to queue message references (Transmit Control Blocks) in the output line processors and handle single message transfers to disc (rather than groups of messages). This has enabled efficient throughput to be achieved using standard moving head disc technology.

Figure 5 shows the data flow associated with an AFTN broadcast type message or an interactive transaction with the local or a remote database. A fundamental data flow aspect is the data interchange maintained between the various 990/10 mini-computers in the system. Two separate memory areas in each line processor are reserved for Cyclic Control Queues. These provide the means of passing high priority control information in each direction between the line processor and central processor. Each control queue is initialised to all zeros (empty state). Two pointers are initialised to point to the same word. These are the LOAD POINTER, indicating the next word to be written, and the ACTION POINTER, indicating the next word to be read. When a new queue entry is added, the LOAD POINTER is incremented, and the alternate processor is interrupted. The queue is then read from the location indicated by the ACTION POINTER. Each word is cleared after it has been analysed and transferred to an appropriate system task queue. Reading continues until a zero entry is encountered. The queue sizes are scaled to avoid over-run (a catastrophic malfunction) under worst case conditions.

Software on the Dataflash has been written in the 990 assembly language, and developed on the development system that will remain in Transport Australia's central office for further design and expansion of network facilities. The central processor uses a commercial operating system known as DX10, which has been developed for the 990/10 computers to control concurrent execution of multiple tasks in a real time environment. DX10 permits efficient task scheduling, interrupt handling, I/O processing and management of files on disc. The line processors themselves schedule their communication functions and internal resource allocation under control of a small com-
commercially available operating system, known as TX990. This runs completely independently of the central operating system.

At initialisation a portion of the memory of each line processor is mapped as part of the memory of the central processor, running under DX10. The initialisation program also downloads the line processor memory, and each line processor then relocates its own memory. The requirement is for the central processor to have access to the Cyclic Control Queues and free chain areas of each line processor. The line processors do not have access to similar areas in the central processor, which acts as a master in the configuration.

In the packet mode, packet transit through the centre may or may not involve data interchange with the central processor, in which case a standard transaction type is identified. The simplified diagram of Figure 5 does not show the many supervisory facilities associated with traffic flow. After initial format check in the input routine, message data is examined at address process time and has to pass various checks. If this input check fails a rejection code is returned to the traffic entry point. Certain other conditions result in an entry to a Job Q routine, for supervisory action. Successful traffic entry results in a header and system number being returned. The address stripping method is used in the network, whereby user addresses are removed (stripped) if further relay responsibility does not exist. Supervisory routines are provided for address table entry/ modification, including group address. The Dataflash software is being structured to enable fully automatic recovery after dual suite failure. After total failure, a hardware-triggered bootstrap firstly dumps memory to disc (crash record) for later analysis. The main program is then reloaded from disc, and the line processors downloaded. Return to normal operation is by means of a background journal recovery program, which references a secure “window” of non-complete transactions in the system, enabling retransmission from the MCBs stored on disc.

11. TERMINAL CONTROLLER SOFTWARE

In the case of the terminal controllers, a small dedicated operating system has been written for maximum efficiency and minimum memory requirements. DX10 has been used for program development, and programming has been at assembly level with the aid of the conditional macro assembler facilities. The operating system designed for the Terminal Controllers is a file-oriented system and occupies approximately 2.5K words. Macros are used exclusively to interface the user to the system. The principle system macro is the file handler called SYSFIL. The general macro library contains an assortment of routines such as nested loops, linear and binary queue handlers, bit maps for logical channel assignments, interrupt driven delays, etc. The macro-approach simplifies writing of source programs and tends to standardise program format. The terminal controller has similar link and network procedures and device service routines (DSRs) to those at the centre. The custom designed operating system has enabled functions that at the centre are implemented over several processors to be efficiently handled in one device. Associated with the operating system is a debug package that may be loaded as necessary. The major software modules and their interaction with the operating system are shown in Figure 6.

Figure 6. Terminal controller software modules
12. TRAINING AND MAINTENANCE

As with other major projects in Transport Australia, the continued satisfactory performance of field installations is of paramount importance. The Dataflash network has a broad range of maintenance requirements, with hardware and software considerations in an operational environment. It can be appreciated that the quantity of equipment involved represents a major increase of programmed systems in the field. The Western Australian network alone has approximately 100 TI 9900 microprocessors. After a review of maintenance philosophy two senior technical staff attended the Austin training facility of TI to undertake formal hardware and software training courses. On their return they were responsible for dissemination of product information and planning and preparation of training material. A comprehensive training syllabus has been prepared covering all aspects of system hardware and software. The first course will be conducted during 1981.

13. CONCLUSION

Emerging data interchange standards are now having a major impact on aeronautical network design. In civil aviation major advances are under way in provisioning of packet switching on the AFTN. For the first time other networks may be readily interfaced and an integrated approach to aeronautical data interchange will ensure significant improvement in the availability and value of aeronautical information. Greater application of data processing in the Air Traffic Service will be evident in the future. Australia is in the forefront of states reaping the benefits to be gained from the new technology.

14. ACKNOWLEDGEMENTS

The author wishes to thank his colleagues on the Dataflash project team for their helpful comments on the preparation of this article.

BIOGRAPHICAL NOTE

W.T. Maloney was educated at Melbourne University, completing BSc (Hons.) in Electrical Engineering in 1960. He joined the then Department of Civil Aviation in 1961 as an Airways Engineer, initially in the VHF Communication and then the HF Communication Sub-Section. In 1973 he was promoted to Principal Engineer in charge of message switching systems. He was elected Chairman of the ICAO Automated Data Interchange Systems Panel in 1977. Mr Maloney is a member of the Institution of Engineers (Aust.) and senior member of the Institution of Radio and Electronic Engineers.
Introduction of a Statewide Computerised Network

A.J. Windross*

This paper examines the planning for a New South Wales State-wide computerised communications network and its implementation. The network is used for the support of more than 2000 cash selling ticket terminals. Particular emphasis is paid to the network outside the Sydney metropolitan area. The substantial implementation problems encountered are described and their solution outlined.

THE APPLICATION

The Totalizator Agency Board of New South Wales (TAB of NSW) conducts off-course totalizator betting on racing events within Australia, and also occasionally on events outside Australia.

The objective identified in the TAB of NSW's Corporate Plan is to continue to conduct off-course totalizator betting within the State, in an efficient manner aimed at providing service to the public of a high quality, minimising costs as far as practicable and maximising the amounts available for distribution to the racing organisations as profit.

Betting operations of the TAB of NSW commenced on 9 December 1964, when six cash selling offices opened using an electro/mechanical, manual and telephone system.

INTRODUCTION OF COMPUTERS – FIRST NETWORK

Cash selling on computerised terminals commenced in November 1971, using keyboard/printers with “NIXIE” displays connected to the host computer on 300 bps half-duplex double-current Telecom line transmitting at 150 bps. Signalling was asynchronous using BCD code. The majority of the 890 terminals in use in 270 offices were connected to the host on copper pairs to a maximum distance of 24 km from the Ultimo (Sydney) Computer Centre. However, some were modified to use Telecom Plan-39 modems on voice frequency carrier to a maximum distance of 60 km.

SECOND NETWORK

A subsequent development was the second network of NSW TAB called TABMARK 1 which is in operation outside the Sydney greater metropolitan area.

The network uses mark sense reader/printer terminals which are installed in cash selling offices of the TAB of NSW. Each office has one to four terminals which are linked by coaxial cable to an office terminal controller. This terminal controller is joined to a Telecom Plan 32 modem for use on a four wire half duplex 1200 bps circuit which may be multi-dropped. At a regional centre each eight circuits meet on a multiplexer. The multiplexer is connected to a Telecom Plan 36 modem on a four wire full duplex 9600 bps circuit to the Sydney TAB head office. There, de-multiplexing to 1200 bps takes place, prior to connection to 1200 bps circuits terminating in a front end communications controller.

The system uses 28, 9600 bps circuits in 12 regional multiplexer centres which in turn serve 224, 1200 bps circuits and 275 offices with some of these offices being multi-dropped on the one 1200 bps circuit.

DEVELOPMENT OF NETWORK DESIGN

By 1974 the TAB of NSW had decided to computerise selling in the outer areas of Sydney and to examine the feasibility of extending this network to the Newcastle and Wollongong areas. Cost justification studies on these plans looked at the major expenditure components such as capital (seven-year write off) and estimated computerised operating cost versus the known operating cost of the existing system. Although the main component of existing expenses was labour, no attempt was made to load the cost to allow for future wage increases.

Early in 1975 the Mark sense reader/printer terminal was ordered for use in the outer Sydney area after the studies which had justified the venture on existing costs and available lines. Parameters included local 1200 bps circuits and Plan 32 (high density) modems. In July 1975, a study showed that a non-metropolitan network using the terminals connected to regional centres with multiplexers at 4800 bps would be marginal with existing costs. Thus, the decision was to defer pending further studies. The month of October 1975, saw a further study based on the non-metropolitan network using IBM type 3704 remote Communications Controllers and the then recently announced Telecom 9600 bps circuits. Some savings on the existing electro/mechanical, manual, telephone system were reported, but before the study was formally concluded IBM had announced the 3705 type Communications Controller. Thus, the study was re-done substituting the 3705 for the 3704. The result was a finding of substantial savings under such a system but a further consideration was introduced with a query as to the amount of hardware engineering support and back-up available outside the Sydney area. Provision of full communications controller redundancy made the proposal unacceptable financially and the matter was again deferred.

REDUCTIONS IN EQUIPMENT COSTS

In February 1976, attention was drawn to price
reductions which had been announced for multiplexing equipment and yet another study was commenced. Still greater savings were calculated to be available in the proposal which also provided for substantial redundancy in remote regional centre multiplexers and at the Sydney TAB of NSW Computer Centre. Telecom was asked to prepare a network plan which took account of the desire to minimise line costs while adhering to the regional centres used in the study. Telecom achieved this within a matter of weeks with only minor line route and regional centre changes. These changes in the main related to availability of plant and trunk route geography. The TAB of NSW then gave Telecom a letter of intent to proceed with the implementation of the network.

However, this advice saw that the network implementation would proceed, as subsequently developed, in phases. These phases being structured as:
- Phase 1 — Outer Sydney area.
- Phase 2 — Newcastle and Wollongong.
- Phase 3 — Other major population centres.
- Phase 4 — Substantial population centres.
- Phase 5 — Other population centres.

The major centre, Broken Hill, had to be excluded because of limited, lengthy and circuitous trunk accesses.

THE EQUIPMENT

The terminal is the IBM type 5934 which was also purchased by the TABs of Western Australia and South Australia. It is designed to operate only on 1200 bps asynchronous lines using BCD code with nine-bit characters. Line control procedures were adapted from the IBM type 2740 Teletypewriter Terminal. The multiplexer, supplied in Australia by Racal, is the T96 Timeplexer. Telecom issued the necessary apparatus attachment approvals (TS72s) in March 1977, and official line orders were then placed.

IMPLEMENTATION

Cutover of Phase 1 (Sydney outer area) commenced on 2 May 1977, but in fact the first cash selling offices converted to TABMARK 1 were those in the city central business district. These 11 offices were selected for a trial of the system because of their close proximity to each other and to the central computer, and because their regular customers were familiar with an automated ticket issuing system.

Conversions in the Sydney outer area followed quickly and by November 1977, the first cutovers were taking place in the Newcastle district.

NEWCASTLE

The Newcastle conversions were approached with confidence as a result of the successful launch in the Sydney city and outer areas but this confidence was soon dispelled. Two particular factors gave cause for concern: (i) complaints of slow ticket issue or system response time, and (ii) delays in line restorations by Telecom following outages. The first problem was met with initial disbelief by the system planners who had recently seen their design calculations apparently proven in practice in the Sydney area. The calculations were re-worked and compared with the results being achieved in the field. A discrepancy remained and this was found to be caused by the amount of data on-line exceeding that expected. A further study revealed the source of this data increase. In April 1977, the TAB of NSW had introduced a new bet type — the Trifecta. The public quickly accepted the Trifecta and by November 1977, it was achieving around 20 per cent of all sales. The Trifecta requires picking the first three finishers in the correct order and thus involves a small but significant increase in data on-line when compared with the Daily Double bet type which the Trifecta had displaced to some extent.

The study also showed that the slowest response times were occurring in the multi-dropped offices. As planned the system had been implemented with two wire half duplex circuits for the 1200 bps legs and four wire full duplex for the 9600 bps regional centre links. The Sydney area offices had no need for multiplexing or 9600 bps as the short line distances did not involve large costs. This lack of multiplexers gave the clue to the performance differences between Sydney and Newcastle. While the multiplexers introduced only minimal delays into the network, these delays were compounded by their interaction with modem turnaround and the polling techniques of the IBM type 3705 Communications Controllers.

FOUR WIRE WORKING

Two wire 1200 bps circuits had been selected because four wire working cost twice as much as two wire for line rentals and to double line rentals would have placed the overall financial viability of the network in doubt. However, the response time problem had to be solved as the service to the public was being degraded with long queues forming on each Saturday morning. When consideration was being given to the feasibility of paying the additional costs of working four wire Telecom unexpectedly announced a major revision of tariffs. Under the new scale, to operate from 1 July 1978, the margin for four wire working was reduced to approximately 30 per cent more than two wire working. The resulting charges again justified the cost of the project and in February 1978, orders were placed with Telecom to introduc four wire half duplex working. Telecom’s response was excellent, and the entire Newcastle region was placed on four wire working by June 1978. Around the same time, Telecom again announced a revision of tariffs. This change removed all four wire rental loadings and the project was then again back to its original cost structure. Without doubt, this tariff revision by Telecom had a major impact on the decision to continue to expand on-line selling through the State.

Four wire working and associated revisions of the polling techniques of the IBM type 3705 Communications Controller had the desired effect and complaints of poor service diminished rapidly. Some problems did remain but these were later redressed by placing on single lines instead of multidrops, certain offices that had improved their ticket issues and sales since the first completion of a network plan.

It is understood that Telecom corrected the second complaint of the Newcastle area, that of slow restoration following outage, by:

(a) Stepping up staff training,
(b) setting up a Telegraphs and Data (Datel) section in the region, and
(c) revising their methods of allocation of data circuits.

Essentially the last point appeared to involve better testing and tuning of the circuit before it was made available to the customer.
Convocations to on-line in the Wollongong area commenced in February 1978, after the decision to alter the network to full four wire half duplex working.

Cutovers proceeded reasonably well until offices on the nearby Southern Tablelands were brought on-line. These offices, particularly those at Picton and Moss Vale, were plagued with line outages and intermittent faults. These two offices are within a 50 kilometre radius of Wollongong, but Telecom had been unable to provide circuits across country, and as a result the circuit was derived Moss Vale/Picton-Sydney-Wollongong-Multiplexer-Sydney. Thus, the circuit actually ran in excess of 200 kilometres. The lengthy circuit and the numerous exchange points without doubt contributed to the high fault rate, and the TAB of NSW and Telecom reached agreement to remove the problem circuits from the Wollongong Multiplexer and terminate them directly at Sydney as single 1200 bps lines. This change and some alterations to amplifier equipment eventually corrected the situation.

MAJOR POPULATION CENTRES

On-line operations commenced in the major cities and towns of the Central Western area from July 1978, and were followed by cutovers in the Riverina, New England and North Coast areas, in that order. Analyses of ticket issues and sales following each cutover proved the assessments of the original feasibility studies and it was determined to eventually place all cash selling offices on-line.

FURTHER PROBLEMS

While the TAB of NSW was pleased with the overall network, concern was being expressed about certain aspects of its operational performance. These performance problems were specifically (i) a high rate of outage on certain of the 9600 bps circuits, and (ii) an apparent 1200 bps line fault, which came to be known as "dropout" and is a situation where a mark sense ticket was inserted and read, but then stopped in the terminal before printing, suggesting that the data was being lost on the line. Remedial action by the terminal operator occupied several minutes and some offices reported more than 50 such occurrences each Saturday.

The TAB of NSW regarded the Telecom response to complaints about the 9600 bps circuit failure rate as somewhat slow but eventually Telecom assigned a "task force" to examine the matter. The task force found what
State-wide Network

Telecom later described as "problems associated with the transmission of data over carrier type circuits" and commenced corrective action. This work was complete by December 1979. While 9600 bps circuit outages still occur they are at a rate far below that seen in 1978 and 1979.

Staff of the TAB of NSW were unable to isolate the cause of the dropouts and both Telecom and IBM Australia were asked to assist. This assistance was readily given and a joint TAB/IBM/Telecom investigation team was set up. The team visited selected offices throughout the State to monitor the problem. In mid-July 1979, while working at the Casino office, they were able to determine that the problem was associated with signal distortion on start/stop bits. Complementary adjustments were made to the IBM 3705 communications controllers and the multiplexers which fully corrected this problem.

SUBSTANTIAL POPULATION CENTRES

Phase 4 of the network was planned for cutover from mid-1979; however, considerable delays were introduced when Telecom found itself unable to readily supply Plan 32 1200 bps modems. Thus, the last four offices of this phase — Bega, Bombala, Merimbula and Eden were not cutover until April 1980. This brings the network to the configuration shown in Figure 1.

OTHER POPULATION CENTRES

These conversions left only 40 offices not served by on-line terminals and it was determined to include these points in a Phase 5 program which was in progress in December 1980. These offices include centres such as Lightning Ridge, Warijala and Tocumwal. Telecom has reported some possible delays due to plant availability although not necessarily in the centres mentioned above.

THIRD NETWORK

Replacement of the original network, installed in 1971, was commenced in June 1980. This new system called TABMARK 2 uses terminals manufactured by Amalgamated Wireless (Australasia) Pty. Ltd., which are similar in concept to the TABMARK 1-IBM 5934 device. The TABMARK 2 device is connected via a line per office on 1200 bps synchronous half duplex circuits. Installation is planned for the greater Sydney metropolitan, Central Coast and Broken Hill areas. The Broken Hill offices are being served by a single 2400 bps circuit direct from Sydney.

CONCLUSIONS

1. The TABMARK 1 network which "pioneered" on-line real time systems on a NSW State-wide basis was successfully implemented because of a phased approach, and cost justification studies carried out prior to detailed technical planning and purchasing.

2. Telecom was not set up to handle State-wide data networks and "learned" at the same time as the TAB of NSW. The implementation problems encountered were probably a factor in the re-organisation of Telecom which has since taken place.

3. Network designers should be continually exposed to the user end of the network, to ensure that performance is not suffering because of changes in data loads or equipment faults.

4. Close liaison should be maintained with Telecom from the inception of the network design up to and including on-line operations.

REFERENCES


BIOGRAPHICAL NOTE

Allen Windross holds the position of Automation and Research Manager at the Totalizator Agency Board of New South Wales, responsible for the planning, procurement, installation and maintenance of all computer derived services. He entered data processing as a programmer in 1968, having previously held a position as Branch Manager of a cash selling office for the Board.
Network Developments at the W.A. Regional Computing Centre 1965-1980

T.A. Reid*

This paper describes the developments which have taken place at WARCC in data communications over the period 1965-80, culminating in the implementation of a multi-host, heterogeneous packet-switched network, with a gateway to national and overseas networks.

Keywords: data communications, computer networks, packet switching, time-sharing, terminals

CR Categories: 1.2, 3.81, 4.32

1. INTRODUCTION: EARLY DEVELOPMENTS

The developments described in this paper were carried out by the Computing Centre of the University of Western Australia and by its successor, the Western Australian Regional Computing Centre (WARCC). This latter body was formed in 1972 to provide computing services to the University, to other educational institutions and to government departments in WA. This involvement with off-campus bodies is partly responsible for the interest the Centre has had in data communications, and goes back to 1964, when it ordered a Digital Equipment Corporation PDP-6 computer. This was the first time-sharing computer in Australia and the first to be commercially ordered anywhere in the world.

Apart from conventional time-sharing (radical at the time, see Moore, Jarvis and Nicholls, 1966), this computer was used to control a number of experiments on-line; these included a flying-spot scanner in Electrical Engineering, a "direct retrieval computer" in Physiology, and a rat-race in Psychology (Nicholls, 1969). These ventures formed the basis for the implementation of a remote computer link, called a remote central processor (RCP). This software and hardware project enabled the connection, via a serial link over data communications lines, of remote computers to the PDP-6 on its I/O bus. It was a fully bidirectional medium-speed link, and was used to connect additional real-time experiments on campus to the PDP-6; thus, the power of the PDP-6 could be used to control experimental equipment such as a diffractometer and a mass-spectrometer. In 1969, such a connection was made over modem-controlled Telecom lines to a PDP-8 at the Commonwealth Scientific and Industrial Research Organisation (CSIRO) in Floreat Park (three and a half miles away). This had a card reader and line printer attached to it, and became the first remote batch station to operate in Australia.

In 1973, the PDP-6 was replaced by a PDP-10 (KA10 processor). The standard data communications front-end for DEC-10s was a DC10, which attached to the I/O bus. Motivated both by a desire to achieve greater flexibility in the front-end and by a shortage of capital funds, the Centre decided to substitute a PDP-11/10 minicomputer for the DC10. This was interfaced to the DEC-10's I/O using a GP10 interface, and communications lines were attached using a multiplexor designed and built by the Centre.

The software running the DC10 emulator handled two of the above 16-line multiplexors, as well as several standard single line interface boards (DL11). Modem control was provided, which was gradually extended to provide dial-up support and then auto-answer. Some of this software was extremely complex, and difficulties were being experienced with supply of the locally-designed multiplexors, so it was eventually decided to install a standard front-end, which then was the DC76 (based on a PDP-11/40), in addition to the DC10 emulator.

Earlier, however, the decision to use a minicomputer as a front-end had proved invaluable; in 1974, it provided the means of connection to the DC10 emulator of a cluster of terminals situated at the Western Australian Institute of Technology (WAIT) in Bentley (seven miles away). A 4800 bps synchronous full-duplex link was used to connect another PDP-11/10 at the WAIT, to which were attached 16 low speed terminals. A simple-minded approach was taken based on the principle of a time division multiplexor. There was no error correction, and line utilisation was very poor.

As the volume of terminal traffic at the WAIT grew, a more efficient and more flexible data link protocol was devised for this link. This resulted, in 1976, in the implementation of a synchronous link to support the 32 terminals at WAIT, utilising DEC's DDCMP protocol, with full end-to-end control, error recovery, better line utilisation, etc. The experience gained in this project also proved invaluable later (see 4 below).

The Centre's Cyber 72 had been installed in 1972, originally with the intention of taking over all the work from the PDP-6. However, it soon became clear that a better strategy would be to concentrate the batch load on the Cyber (with some on-line file enquiry applications), and to retain a specialised interactive facility. This was what led to the replacement of the PDP-6 by a...
Network Developments at WARCC

Hosts
Transport Mechanism
Terminals
To Other
Other Facilities
To Other Networks

Figure 1. A network concept

PDP-10 in 1973; and our experience has corroborated the views of many that if there is a need to provide good interactive and good batch service, then it will be achieved more effectively (and cheaply) by separate computers than by a single unit. However, it was recognised that an important ingredient in this mix could well be some kind of file link between the two computers, primarily to allow use of the greater power of the Cyber, in batch mode, by DEC-10 users.

Due to manpower constraints, this was not achieved (other than via magtape) until 1977, when a 1200 bps asynchronous link was established over standard data communications interfaces. Each computer considered it had a standard asynchronous terminal at the other end (see Figure 2). The principal components are a program in the DEC-10 which transmits spooled requests for transfers to the Cyber, and a program in the Cyber which receives and processes this flow. This system enables a DEC-10 user to submit a job to any of the Cyber queues (Input, Output, Punch, etc.), or to catalogue a file on the Cyber disks. Thus, a form of "edit and batch submit" was provided; so was the ability to make use of the Cyber printers (either as back-up for the DEC-10 printer, or to make use of the higher speed of the Cyber printers); to make use of the Cyber card punch (one is not provided on the DEC-10); to make use of the Cyber spooler for the off-line plotter (such a spooler has only recently been implemented on the DEC-10); to enable a DEC-10 file to be printed at a Cyber remote batch station; or otherwise to gain access to the wider range of facilities and packages available on the Cyber at the time. An important aspect of this project was, of course, maintenance of account integrity, especially in connection with creating files on the Cyber disk.

Later the same year (1977), another link was established through the DC10 emulator on the DEC-10, this time to an HP21MX. This was essentially a revision of the 1974 remote terminal concentrator, and permitted terminals connected to the HP21MX to connect, through that mini, to the DEC-10 (see Figure 2). The WAIT Links were removed in 1977 when WAIT acquired their own DEC-10.

These activities span a considerable period, and cover a wide variety of functions. They have provided the experience, and in some cases the basis, for the more ambitious developments described below.

2. CYBER EMULATORS

Some of the work described above, and particularly the experience in programming PDP-11s for data communications tasks, suggested several uses directly associated with the Cyber. In particular, given data communications line costs in Australia, and the cost of standard CDC terminals (711, 714), an obvious application was to use a remote PDP-11, connected over synchronous lines to a standard CDC interface, to emulate multi-dropped 711s or 714 clusters. As a bonus, greater flexibility and adaptability would be provided.

Accordingly, in 1975 the first 714 emulator was installed. The immediate rationale was to increase the number of lines on the Cyber, provided by 6671 MUXs. An additional factor was the need to provide support for remote reading, by the Public Works Department, of non-standard paper tapes (containing river-gauging data). The best that can be said for the modifications made to the emulator software to support paper tape is that they were made hastily!

Partly as a consequence of this, it was decided to rewrite the 714 emulator software, and this was accomplished in 1976. On this occasion, great pains were taken to write code which would be easy to maintain and which could be readily modified or adapted. Use was made of...
the Duke University structured macros for the PDP-11 (Herman-Giddens, et al., 1975).

At the same time, the software was adapted to provide remote batch terminal (200UT, 734) emulation on the Cyber. The motivation was again to provide a cheaper alternative to the standard CDC equipment, the emulator being several thousand dollars cheaper to purchase (including software), as well as being nearly half as expensive to maintain. At this time, the 714 emulator was cost-effective if four or more terminals were required (using asynchronous terminals such as the Teleray 3311). The use of a good PDP-11 cross-assembler on the DEC-10 was an essential component of these developments.

Also in 1976, a 714 emulator was implemented on an Interdata 6/16 minicomputer. The availability of software modules on the 6/16 to handle many of the data communications functions (e.g., terminal and line interface support) promised significantly to simplify this exercise. Such, however, did not prove to be the case, with many bugs and deficiencies being revealed in those modules before the emulator was finally installed at Murdoch University.

Some 20 or 30 installations of the various versions of the 714 and 734 emulators have now been made in Western Australia, on all five Cybers in Perth. The decision to follow this approach has been well-justified, as the 734 emulator has been adapted to provide a multi-drop capability. This has allowed installation at the one site, on the one line, of mixes of multiple card readers, line printers and VDUs. It has also proved invaluable to Health Computing Services, with their diverse needs in providing computing services to Perth's teaching hospitals. One of the adaptations pioneered by HCS has been the support of “site poll”, which has dramatically improved response times on emulator VDUs.

3. CONVERSATIONAL REMOTE JOB ENTRY

Experience with the DEC-10 to Cyber link, described in 1. above, led the Centre to believe that the increasing pressure to provide improved interactive access to the Cyber could be largely satisfied by providing an adequate “edit and batch submit” facility. This belief was bolstered by similar developments being undertaken overseas and at Melbourne University. Accordingly, it was decided in 1978 to acquire a PDP-11/60 which would provide a good, cheap editing capability, and to link it to the Cyber using our 734 emulator software.

This was implemented in 1979, with the emulator running in a PDP-11/03, back-ending the 11/60. The batch-submit functions are implemented by jobs which run under the RSTS/E operating system. This makes them much easier to implement (they are written in BASIC), and use can be made of a suitable existing editor, such as SOS. A spooler program empties the queue of jobs being submitted to the Cyber, sending them as files to an asynchronous “terminal” which is the 11/03. Output being returned from the Cyber is held on disk by the CRJE pending action by the user. The user can attach to such files, examine them with the editor, save them on his own disk, purge them or have them printed at a nominated or default printer attached to the CRJE. Alternatively, he can have his job automatically printed at the central site (or on some RJE printer) using the Cyber ROUTE command, or have it automatically printed on a specific printer on the CRJE.

A bonus is the access so provided (for authorised users) to the other facilities of the RSTS/E operating system environment, such as the WORD-11 word processing software now installed on the 11/60. Many of the facilities for support of interactive terminals are much better under RSTS/E on the 11/60 than under NOS/BE (or even NOS!) on the Cyber. The CRJE is an important step in the direction of elimination of cards and card punches, particularly for students, while preserving many of the advantages of a rapid batch service. Instead of acquiring and maintaining card punches to support the increasing number of students making use of the Cyber, we can now substitute an equivalent number of VDUs, at many times smaller cost per unit, and with considerably greater flexibility.

4. PACKET SWITCHED NETWORK

As early as 1974, the Centre began investigating the applicability of packet-switched networks to handle the data communications needs of Centre users, and indeed, of most computer users in the public sector in Western Australia. There had, of course, been some spectacular disappointments in this area, which prompted us to tread very cautiously. Furthermore, there were significant political implications in any such “global” approach, with some quarters openly concerned lest a body to control all public sector computing ride in on the back of a general network.

Nevertheless, aspects of this approach are sufficiently attractive to warrant careful analysis. For example, it would only be by collaborating that most of the various public sector organisations could afford data communications to country towns in Western Australia. Accordingly, a report was prepared in 1977 calling for an in-depth study of the needs for such a network in Western Australia (Reid, 1977). In response to this call, a Western Australian government task force has collected data on traffic volumes, growth patterns, etc., in all public sector organisations in Western Australia. Discussions are continuing with Telecom to determine how soon a suitable service might be offered publicly; at present, it still looks as though 1982-83 will be the earliest that this might happen in Western Australia. Five years is not a very long time when developing facilities such as this, so a high degree of consultation is required.

At the same time as the above discussions were taking place at the political level, the Centre had been investigating the technology of packet-switched networks. Some less ambitious projects with limited objectives had been very successful, e.g., Merit in USA (Scott, 1978) — but even here it had taken a long time for use of the network to take off. Of those projects with broader objectives, only ARPANET had much written about it, and the magnitude of that exercise was well beyond the reach of our resources.

We were casting about for a strategy to provide the following functions:

- rationalise the diversity of communications links existing or envisaged among the equipment in the Centre or elsewhere;
- provide a degree of insulation for terminals from changes in hosts (i.e., limited “virtual terminal” capability);
- provide access by as large a population as possible to a wide range of facilities;
provide access to off-campus (and interstate and overseas) facilities;
— enable terminals on a State-wide basis to gain access to the Centre's and other facilities;
— provide a high degree of flexibility and adaptability in any communications offering;
— and not least, we wanted something that would work!

Overall, we needed a “transport mechanism” which would provide for terminal and file traffic between an array of suppliers (hosts, other networks, etc.) and consumers (terminals, other hosts, minis, etc.); and which would do this in a rational, consistent fashion (see Figure 1).

The offerings of the computer suppliers were canvassed, but little was available here. Control Data offered nothing; most other suppliers offered nothing; IBM had announced SNA, but that seemed to be ruled out on the grounds of extreme expense and dearth of documentation, so the connection of non-IBM hosts looked impossible. Only DEC's offering, DECNET, seemed promising.

Obtaining information about DECNET proved difficult. Nobody in Australia knew anything about it. Eventually, it proved possible to obtain a copy of the software, with very limited documentation, thanks to the efforts of the staff of the local DEC office. Indications were that this product would satisfy most of the above needs. Furthermore, it seemed possible to implement this software between the two DEC-10s at WAIT and at WARCC with almost no effort.

Surprisingly, this indeed proved true, and by spring 1977 a packet-switched, multi-host network was installed between WARCC and WAIT, using DN87 front-ends (upgrading the DC76s with the DECNET software), and a synchronous 4800 bps line (see Figure 3). More surpris-ingly, almost no bugs were found in this code! This is believed to have been the first multi-host packet-switched network to be operational in Australia.

The next stage was to adapt this software to interface to the Centre's Cyber as well. The approach taken was to make the Cyber look like a DEC-10 to the network, and the network like a 714 cluster to the Cyber. The first part of this was achieved by taking the software which ran on the DEC10 to service network traffic, and to rewrite it to run on a PDP-11, calling it “NET11”. Software to emulate 714 clusters on a PDP-11 already existed, but this was rewritten to achieve greater efficiency in PDP-11 CPU and memory, and was reduced in size to a few pages of code.

The three elements of this software were then interfaced and installed on a PDP-11 (designated a “CN87”) in front of the Cyber's 2550 (see Figure 4). A full discussion of the development of this software (including how it was tested) can be found in Fernandez (1978).

This expanded network went into operation at the beginning of 1978. An important aspect of the development had been the principle that no changes should be made to the host operating system (in this case NOS/BE). DECNET employed a three-layer approach, and these layers and their interfaces were strictly defined and policed (DEC, 1976).

It had also been intended to connect a few PDP-11 hosts located in other educational institutions, but this proved more difficult than at first hoped. A discovery was made that DECNET came in two incompatible flavours—a version for DEC-10s (using ANF-10 software), and a version for minis which was DECNET proper. The latter did not have a DEC-10 interface (nor does it have yet), but was to become the main offering of DEC for data communications networks. The former, the version we had employed, was not going to be developed further by DEC.

Accordingly, in order to incorporate a PDP-11 into the network, we had to adopt the same approach as for the Cyber; i.e., we had to make a PDP-11 look like a DEC-10. This did not prove too troublesome, as we already
5. FUTURE NETWORK DIRECTIONS

The above development of a packet-switched network has been carried out on a shoe-string budget, and has been experimental in nature. The results achieved have been very encouraging, though many loose ends need to be tidied up. This approach has proved to be a viable framework within which to provide for an ever-increasing diversity of needs. At present, terminal switching is essentially the only facility provided (and a device such as a private automatic computer exchange [PACX] would be both more efficient and more economical for that function alone), but the following developments will fit well in the framework adopted.

(a) File Transfer

File transfer between hosts will be an important requirement, for many reasons but primarily to allow use of different packages on different computers. Although magnetic tapes still represent the fastest data communications medium, the trend is decidedly away from them: for all but very large files they are much slower than use of a network; considerable incompatibilities still exist between tape files on different computers (even from the same supplier); it is an operator-intensive activity; it relies on use of mechanical devices; and in general is a labour-intensive activity (add up how much time is spent untangling tape transfer problems between computers!).

File transfer from minis to large computers will be (and already is) important to enable analysis by the mainframe, perhaps using specialised software packages, of data collected by the mini; to allow access to the number-crunching or data processing capability of the mainframe; to enable incorporation into a central data base of data collected at diverse local facilities or just to process regional data against a central data base; to make local results available to a wider population; to enable batch submission to take place from decentralised equipment; or otherwise to make accessible the specialised facilities provided on the larger computers.

File transfer from the mainframe to mini is needed to support cross-assembling and down-line-loading functions, to return results of mainframe processing to the originating mini, to support remote printing, to provide for rationalized support for graphics, with some processing being done on the mainframe, but high-speed interactive support provided at the mini, etc.

(b) Connections to a wider variety of Equipment

The variety of equipment and services is continually increasing, and access by workers in educational and research institutions (in particular) is vital. In this category would be access to the wide variety of packages available on IBM equipment (there are currently almost no IBM computers in educational and research institutions throughout Australia). There is going to be a need to gain access to other specialised facilities, such as array processors, data base processors, information retrieval systems, etc. Australia as a nation cannot afford, for example, much more than one computer in the class of the Cyber 203, and if its researchers are not to be at a disadvantage compared with colleagues in the UK or USA, then they must be given access to such facilities via networks.

(c) Connection to other Networks

Part of the above variety of services will be provided if rationalised links can be provided to existing national and overseas such as CSIRONET, Tymnet, Tele-

Network Developments at WARCC

net, ARPANET, etc. This is best provided by means of gateways, which ultimately will be facilitated by widespread use of standards in the data communications area such as X25.

Furthermore, such links can foster a high degree of communication between researchers, nationally and internationally, and between researchers and government and industry; this is particularly relevant for Australia, with its small, widespread and isolated population. Much research thrives on consultation, collaboration and co-operation; the fuel crisis is making it too expensive to have much face-to-face contact, and Australia will be at a particular disadvantage. Data communications may provide a way of compensating for this, and ensuring that we do not fall behind or fail to attract and retain competent researchers.

(d) Electronic Mail

Already, experience overseas (e.g., in ARPANET) suggested that electronic mail is going to be one of the most important functions provided by networks (see, for example, Uhlig, Farber and Bair, 1979). Researchers, particularly, feel very isolated in Australia, and electronic mail holds out the following advantages over conventional communications media:

- easy, rapid transmission of hard copy;
- broadcasting of messages to groups of colleagues;
- avoids the frustration of the telephone; and
- avoids the delays and uncertainty of the mail systems.

6. CONCLUSION

Considerable difficulties have yet to be overcome before we reap all the benefits of 5 above. It is clear that the present attempts at a network at WARCC will probably never provide all these benefits, and incur several frustrations for users. There are substantial overheads in the use of packet-switching, especially when dealing with a DEC-10 in full-duplex mode over the network. Furthermore, it is by no means clear that the costs of a network are justified for the present volume and style of traffic, compared with other mechanisms such as a PACX for local traffic, or the switched telephone network for off-site traffic.

On the other hand, the cost of data communications is dropping dramatically — perhaps not as fast as the cost of computer hardware, and perhaps not so fast in Australia as elsewhere because of the high degree of regulation. But the use of general purpose networks such as that pioneered at WARCC is here to stay. To oppose this trend would be like trying to save the Pony Express, which went bankrupt two years after Western Union introduced the telegraph to the West.

Only such a general-purpose network can hope to provide the full range of facilities needed, or provide the flexibility to adapt to changing demands and changing services. The launching of a nation-wide packet-switched service in December 1982 will be an important step for Australia. However, one can be forgiven for fearing that the cost may deter much ad hoc use, at least in the first instance, unless a clear vision of the benefits is retained. This is nowhere more true than among the various research institutions in Australia, which have to date been very slow to realise the importance of access to a national network. This is not so in USA where Peter Denning, ACM President, in reporting on a meeting between the Council of Scientific Society Presidents and the office of the President of the USA, notes: "A step of high potential impact is support of national computer networks, which would permit exchanging documents, programs, and data among scientists; networks would also permit remote use of specialised equipment" (Denning, 1980).

Providing the raw communications channels, albeit packet-switched is, however, only part of the answer. Access to the range of facilities and services available on any national network needs to be "facilitated", especially if they are to be used by non-computer professionals. In the USA a body has been set up specifically to accomplish this (Heller, 1978). This may not be the pattern for Australia, but a real need will exist and will have to be addressed.

7. ACKNOWLEDGEMENTS

The work described in this paper has been performed by only a handful of people, who have demonstrated a high degree of dedication and hard work to see these developments realised. Chief among these are Terry Gent, Bruce Kirkby, Joe Fernandez and Alan King, to whom WARCC is very grateful for their enthusiasm and effort. The author also wishes to thank them for the comments they have passed on this paper.

REFERENCES


BIOPGRAPHICAL NOTE

The author entered computing in 1961 at the Weapons Research Establishment in South Australia, and then worked for the Bureau of Statistics before joining the staff of the Computing Centre at the University of Western Australia in 1969. He became Director of that Centre (now the Western Australian Regional Computing Centre) in 1979. He has a BSc degree and the Dip Comp from the University, and spent three years lecturing in the Diploma in Computation program there. He was awarded the ACS Case Study Prize in 1978 for his paper, "The Trials and Tribulations of an On-Line Computer Project".
1. INTRODUCTION

The Sydney Unix Network is a network of computers within the Universities of Sydney and New South Wales, that operate under the Unix Timesharing System (Thompson and Ritchie, 1974). The development of the network began in 1978 with the installation of a leased line connecting a PDP11/40 computer in the Basser Department of Computer Science at the University of Sydney (known as "basser40") and a PDP11/70 at the Australian Graduate School of Management on the University of New South Wales campus (known as "agsm"). This connection was used initially simply for access to the AGSM machine from a terminal connected to the Basser 11/40. Subsequently within each campus a number of other links were established and used in a similar way, i.e., a terminal on one machine could be used to login to a remote machine. Users wishing to login to a remote machine would invoke a special program that copied characters between their terminal and the outgoing line. At the remote machine the incoming link appeared as a normal terminal connection.

The limitations of this initial system were: (1) only one user could use the link at a time even though the available bandwidth would satisfy many users; and (2) the line was directional. This may be explained as follows: each computer running under Unix associates a simple login program with each incoming line. This program listens for activity on the line and then normally engages in a login dialogue with a remote human user. If a line connecting two computers is regarded as an incoming line by both, then the two login programs may engage in an endless and fruitless dialogue. Worse, if a connection is interfered with the subsequent activity on the line. The protocol allowed the lines between machines to be replicated very early in the network development as a solution to the problems that arose with the original single user, directional terminal connections. The multiplexing protocol allowed the lines between machines to be replicated by software. Thus, more than one user could be using a physical line at the same time and several virtual lines could be designated as "to" a remote machine and several as "from" the same machine.

In early 1979, wider issues of network development began to be considered. These included file and mail transfer systems and virtual circuit facilities. Conventional ways of providing these services involve building a network of node computers connected together by reasonably fast lines and with host machines connected to the network node computers. This solution was not seriously considered since in many cases the cost of a node machine would be comparable with the cost of the host it was to handle. Instead a host supported network system was developed that used no specialised network node computers and utilized the fixed and leased communication lines that were previously used as simple terminal connections.

2. PROTOCOLS

The protocols used in the networks do not, at present, conform to a strict layering as in other more conventional networks. For example, the lowest level of software protocol used over the physical communication lines is not strictly necessary for the higher levels, although it is used for all present links. This will be made clear in the description of the higher levels of protocol.

2.1 Physical Link Layer

The first "layer" of network protocol consists of the physical lines. These are either fixed lines running between machines in the same building or on the same campus or leased from Telecom. The interface to each machine conforms with the RS232 standard. The line speeds range from 1200 baud for the leased lines between campuses to 9600 baud for the lines between machines in the same building. The current topology of the network (February 1981) is shown in Figure 1.

2.2 Multiplexing Layer

The second layer is the lowest level of software protocol. It involves a simple multiplexing scheme that allows one physical line to appear to the operating system as several virtual lines. The protocol used in the networks do not, at present, conform to a strict layering as in other more conventional networks. For example, the lowest level of software protocol used over the physical communication lines is not strictly necessary for the higher levels, although it is used for all present links. This will be made clear in the description of the higher levels of protocol.

The multiplexing protocol allows the lines between machines to be replicated by software. Thus, more than one user could be using a physical line at the same time and several virtual lines could be designated as "to" a remote machine and several as "from" the same machine.

The multiplexing line driver and protocol are known simply as "mx". There is very little in the way of error control. However, flow control is provided.

The mx protocol multiplexes "ports" onto real lines and simply separates data for each port on the line. Each block of data has a four byte header:
This header is easily recognised and checked for consistency. The length of the data block that follows the header is usually limited to 40 bytes. Flow control is implemented by sending blocks without any data, the first byte count zero and the second byte count representing the number of data bytes that may be received. The sender must not send more than the specified number of data bytes before waiting for another flow control block. A simple timeout mechanism is used to recover from lost flow control blocks.

2.3 Transport Layer

The next network protocol layer may either be a virtual circuit facility or file transfer. The virtual circuit scheme is simple but effective and grew out of the original machine-machine connections that allowed a user to connect his terminal "through" his local machine to a remote machine. (See Figure 2).

The program now used to make the connection to the remote machine is called con (there have been other similar programs designed for special situations). This program takes as argument the name of the remote machine and from this determines the correct outgoing link to use. The program then performs a system call that would instruct the line driver to "connect" the terminal and the outgoing line. Characters are transferred from one line to the other very efficiently. The outgoing link may in fact be a multiplexed link created by "mx", along with several others, from a single physical line. This is hidden from the user and the con program.

In order to allow a user to connect to a remote machine that is further away, i.e., that involves more than one intermediate machine, the con program carries out the extra function of routing. As part of the file transfer system, a network topology file is maintained. The details of this file will be described later. However, it does contain a list of all the network hosts and machines they are directly connected to. For example, from Figure 1 the topology file (called "netstate") would contain an entry for the host "basser40" as follows:

```
basser40 -* agsm
   -* chemeng
```

and the entry for "basservax" is:

```
basservax -*■ basser40
   -* elecvax
```

The con program consults this file before making a connection and determines the route of such a connection. For example, if a user at a terminal that is directly connected to the "basser40" types "con elecvax", then the con program consults the "netstate" file and discovers that the route from basser40 to elecvax is via basservax. It then sets up a connection to basservax and sends a special message to the login process at basservax that indicates that a connection is required to elecvax. The login process at basservax then in turn starts a con program with "elecvax" as argument and the process repeats. When the login process at elecvax receives the special message (now generated by a con program on the basservax) it ignores it and the login proceeds as usual.

When the user has finished his session and logged off the remote machine, he types a special escape character that is interpreted by the con program running at
the local machine (A in Figure 3). This machine will in turn pass an escape character on to the next machine in the route (B in Figure 3) and so on until it reaches the remote machine. The con programs running at the remote and intermediate machines terminate, returning control to the login process in each case. The con program at the local machine also terminates and the users command interpreter regains control.

This scheme is very simple and has proved to be an effective solution. Connections are made from one side of the network to the other without problem.

3. FILE TRANSFER LAYERS

The other major part of the network is the file transfer system. It is implemented in three levels that may or may not be used above the "mx" level. It is possible for it to operate on dedicated lines between machines but in practice it uses one port of an mx controlled line.

The top level is the user interface layer, the second level, implements an end-to-end protocol and maintains the network topology file, while the lowest level provides an error-free path between directly connected nodes. Figure 4 is a diagram of the protocol levels of the system.

3.1 User Level

At the user interface level, files, electronic mail or printer output may be transferred by a user from his home machine to any other on the net by using a single command. Network addresses are of the form "username:hostname" where "username" is a valid login name on the remote host whose network name is "hostname". This form of address is recognised by the mail program and the mail item is then handled by the "net" layer of the protocol.

The file transfer system is implemented with two user commands. These are "netsend" to send one or more files to a remote host, and "netget" to retrieve a file at a user's local host which has arrived as the result of a netsend on a remote machine. The netsend command has at least a username:hostname argument and may have one or more file name arguments specifying the files to be...
Figure 4. End to End Level.

sent. If no file name is specified, the standard input file is used. This is usually the users’ terminal but may be a “pipe” from another program. The netget command takes zero or more option and filename arguments; when used without arguments it retrieves files that have been sent from other users or hosts and placed in a holding directory (or “spool”). It is important to note that netget does not allow a user to reach out from his local machine to some remote host and copy files. It is one of the major design decisions of this network file transfer system that this facility not be provided, the main reasons being security and the desire for simplicity. If the facility were available, a system of network wide user validation would have to be implemented. Since almost all the machines on the net are independently managed, and some have large numbers of potentially hostile student users, it was felt that such a scheme would be difficult to implement in a secure way. With the current system a user wishing to get a file from a remote machine would invariably have a valid user-name on the remote machine and so would be able to use con to the remote machine and “netsend” the file to his local machine.

The other service currently available at this level is a remote printing facility. A user may use the command “netprint remote-host-name file-name” to print a file on a printer attached to a remote host. The file is sent using the normal file transfer system but is passed to the line printer handling program (or “daemon”) when it arrives at the remote host.

Also available to the user are two ancillary programs, “netq” and “netstate”. The “netq” program gives information about files waiting to be sent to directly connected hosts. The origin, destination, size and position in the relevant queue are given. The “netstate” program lists the network topology file in a readable form.

3.2 End-to-End Level

The next lower level, implemented by the program “net” (see Figure 4), accepts files for transmission, determines the next node in the route to the destination host and saves the file together with an end-to-end protocol header in a directory set aside for the next host in the route.

The end-to-end protocol header contains the following information:

- destination host name
- destination user name
- origin host name
- origin user name
- action
- file name
- action parameters
- statistics

When a file is received by the lower “netd” level it is passed to the net program for processing. If it is destined for the local machine it is placed in a directory to be retrieved eventually by a user via the netget program. Also a message is sent (by system mail) to the user notifying him of the arrival of the file. A time limit is placed on the files in the reception directory after which they are
removed. The user is warned of this time limit in the arrival notification message. If a received file is enroute to another host, it is placed in the appropriate directory and the "netd" level program will send it on to the next node in the route. The next node is calculated by the "net" program from information found in the "netstate" file. This is a "staging" system that involves complete reception of a file at an intermediate machine before sending it on to the next machine in the route. This has the disadvantage of requiring enough file storage at intermediate nodes to contain all the in-transit files, but this has not proved to be a problem. The benefit of the scheme is that it greatly simplifies the end-to-end protocol and allows routing decisions to be distributed across the network.

The net program also maintains the network topology file called "netstate" which contains an entry for each host. Each entry includes the time that the named host was last active and the transmission rate achieved. Following this is a list of hosts that are directly connected to the named host. Example entries have been given earlier. Each time the lower level "netd" program is started it signals "net" to send a message to all directly connected hosts indicating that the local host is now active. The directly connected hosts propagate this "host-up" message through the rest of the network and also send a copy of their "netstate" file back to the new host. When a directly connected host does not respond to some message and is deemed to be down, the netstate file is modified to indicate the new state. Thus, the network topology information is maintained by the "net" program to reflect the current state of the network.

3.3 Node-to-Node Level

Below the "net" program level the node-node file transfers are handled by a program known as "netd" or network daemon. In each machine there is a copy of this program executing for each directly connected host. For example, in the basservax host there are two invocations of netd, one for the basser40 and one for the elecvax host. In the elecvax and baser40 there would be a netd program running to handle the link to basservax. The netd programs at each end of a given link co-operate to transfer files from one node to the other.

Netd uses a half-duplex, multi-buffered, positive acknowledgement data transfer protocol. Before each file transfer the netd programs negotiate the direction; file transfer then proceeds with short data blocks enclosed in a protocol envelope.

The header consists of the following 8 bit bytes:

- SOH (ASCII start of header character)
- sequence number
- size of block
- complement of size of block

while the trailer is:

- longitudinal parity check
- sum check

The sequence number is used to provide the multi-buffered flow control. Each packet must be responded to with a two byte reply consisting of an ASCII ACK or NAK character and the relevant sequence number. Errors cause the retransmission of all un-ACKnowledged packets. Catastrophic error conditions cause a negotiation for restart of the file transmission. The error check characters were chosen as the simplest to compute and yet have a reasonable chance of detecting errors.

The netd program executing on behalf of a particular host-host link scans a designated file directory for files to be sent. The files are placed in the directory by the "net" program.

4. IMPLEMENTATION EXPERIENCE

The system was implemented by one of the authors (PRDL) over a period of two months with a total time spent on the project of six man weeks. It has now been operating for six months and has fulfilled the original aims of allowing closer co-operation between system maintainers and developers. Other, non-system programmer users are now beginning to use the system for such things as preparation of teaching materials.

The programs are all written in the programming language C (Kernighan and Ritchie, 1978) except for the "netsend" and "netprint" commands which are command interpreter procedures that invoke the "net" program with appropriate parameters. The number of lines of C code is just under 7000. The size of the binary code of "net" and "netd" is approximately 22,000 bytes and of "netd" is 19,000 bytes. The system has been implemented on a number of different machines and versions of Unix. It is currently operational on Vax 11/780, PDP11/40, PDP11/34, PDP11/60 and PDP11/70 machines under Unix versions 6 and 7.

5. FURTHER WORK

While the file/mail transfer system is operational only on systems running the Unix operating system (but on a variety of different machines), there are a number of non-Unix machines that can be reached via the con program. These include CDC Cyber machines at both Sydney and NSW University computing facilities. Recently a link has been established between the baser40 and a CSIRONET node machine located on the Sydney University campus. This link uses the "mux" multiplex protocol and allows several users simultaneous access to CSIRONET hosts. It also allows users connected to the CSIRONET to connect to the Sydney Unix Net.

The network will undoubtedly grow with more Unix hosts (at least three more are expected at Sydney University in 1981). The main enhancements to the network software will involve gateways into other networks such as CSIRONET and experiments with trans-network file and mail transfer. This will entail modifying or adding to the existing programs and protocol layers. For example, a trans-network mail system that uses the facilities of CSIRONET and the American Telenet is planned. This will allow mail transfer between local users and colleagues in America.

6. CONCLUSIONS

The Sydney Unix Network is a simple but effective system that provides a virtual circuit facility, a terminal connection system and a file/mail transfer system. The simplicity of the design meant that a comparatively small amount of effort was required to implement the system and it has proved to be very easy to maintain. At the user level it provides an effective means of co-operation between users at widely separated sites.

7. ACKNOWLEDGEMENTS

The design of the Sydney Unix Network software was influenced by many people. System managers at most of the host machines contributed ideas and suggestions. Especially helpful were Peter Ivanov, Kevin Hill, Andrew Hume, Chris Maltby, Chris Rowles and David Horsfall. Ian Johnston contributed to the early development of the system. The link to the CSIRONET was made possible through the efforts of John Gibbons.

8. REFERENCES


BIOGRAPHICAL NOTE

Robert J. Kummerfeld received a BSc with first class honours in Computer Science from the University of Sydney in 1973 and a PhD in Computer Science from the University of Sydney in 1979. Dr Kummerfeld has been a lecturer in the Basser Department of Computer Science at Sydney University since January 1978. His research interests are computer networks, distributed processing and interactive systems.

Piers Lauder was born in Nicosia, Cyprus. He received a BSc in Engineering and Economics from Warwick University, UK in 1969, and a Postgraduate Diploma in Computer Science from Bradford University in 1973. Since 1974 he has been a programmer with the Department of Computer Science at Sydney University. His major interest is in networks.
Decentralised Scheduling

L.M. Casey*

Multicomputer systems, including networks, that are going to realise their full potential for fault tolerance combined with cost effectiveness will need to employ decentralised control. In particular scheduling decisions will have to be decentralised decisions. After comparing decentralised scheduling with its alternatives, this paper goes on to survey mechanisms that have been proposed for decentralised scheduling. Details (including some simulation results) are given for a distributed load balancing scheme that employs latest reported status. This scheme minimises bandwidth requirements, while evenly spreading the work load.

Keywords and phrases: Load balancing, scheduling, dynamic, distributed, decentralised.

CR Categories: 4.32

1. SCHEDULING IN MULTICOMPUTER SYSTEMS

Multicomputer systems, including networks, that are going to realise their full potential for fault tolerance combined with cost effectiveness will need to employ decentralised control. In particular scheduling decisions will have to be decentralised decisions. Scheduling is the determination of the order in which competitive claims for resources are met. For a computation to progress it needs processor time and memory space. It may also require access to peripheral devices. Normally in a multicomputer system these resources have to be co-located on the same processing element (PE).

Scheduling can occur at several levels within a system (see Figure 1). There can be local scheduling. This is the management of resources local to a PE by the kernel/executive of the PE. Thus, when there are a number of tasks loaded and ready to run on a PE it is that PE's decision which to run first. Local schedulers are not concerned with the effect their decisions have on the total system.

Global scheduling is basically the determination of which PE will execute what unit of work. As well as the availability of processing power at a PE the scheduling decision has to take into account the availability of otherwise of memory and any required peripherals. The global scheduling process may also take account of the bandwidth requirements of moving a task from one PE to another. Another aspect might be the determination of the current number of replications of (the code for) a service or function in the multicomputer system.

1.1 Static and Dynamic

Global schedulers can be static or dynamic. Static scheduling is the a priori assignment of tasks to PEs. It may also involve fixing the order of execution of tasks, if this function is not carried out by local schedulers. Dynamic scheduling is run time scheduling. Although making decisions at run time involves extra computation, dynamic systems have a superior potential for availability, reliability and extendibility. This superiority arises from the nature of dynamic scheduling which is to adapt to changed circumstances, including changes in system configuration.

As well as enhanced system availability and extendibility, dynamic schedulers may give a superior system performance compared with static schedulers. A lot of work has been done on the static assignment of tasks to PEs when a system has a completely characterised and totally constant workload. For a system with fixed hardware resources it is sometimes possible to determine a placement and order of execution of tasks that is an optimum (according to some criterion such as minimum completion time). Frequently, determination of the optimum is a very complex computational task. Much work has concentrated on finding heuristic methods for computing near optimal solutions (see the bibliography of Satyanarayanan, 1980).

When the system's workload is not constant or resource requirements are not totally known in advance then the static assignment of functions to PEs involves a performance penalty. Variations in the work load will result in some PEs becoming overloaded while others are idle. A dynamic scheduler adjusts to fluctuating workloads. Depending on the mechanisms employed, the extra overheads involved in dynamic global scheduling should not totally negate the performance margin it has over static assignment. Thus, smaller systems would be able to carry out the same work.

1.2 Centralised and Decentralised

Again dynamic global schedulers can be one of two types: centralised or decentralised. They are centralised if there is a single decision point and decentralised if decision-making is spread throughout the system.

Centralised scheduling does have the advantages of simple implementation and the best theoretical perform-

Figure 1. Classification of schedulers

*Scheduled *Copyright © 1981, Australian Computer Society Inc. General permission to republish, but not for profit, all or part of this material is granted; provided that ACJ's copyright notice is given and that reference is made to the publication, its date of issue, and to the fact that reprinting privileges were granted by permission of the Australian Computer Society.*

*The author is with the Computer Research Section, Physics and Engineering Laboratory, DSIR, Private Bag, Lower Hutt, New Zealand. Manuscript received 27 March 1981.

Decentralised Scheduling

2. SURVEY OF PROPOSED LOAD BALANCING MECHANISMS

To optimise performance a decentralised scheduler should try to minimise the amount of time that a PE in the system is idle while there is work queued up at another. This is the intuitive concept of load balancing. It is difficult though to produce a quantitative measure of load balancing. In the survey below we shall not attempt to quantify the load balancing effectiveness of each type of scheduler. Rather, we examine such features as stability and communications overhead. It should be noted that the impact of a scheme that has high message traffic may well be on processing power rather than bandwidth. To process a message, a processor has to interrupt its current activity, deal with the message, and then resume the interrupted activity.

One way to achieve decentralised scheduling would be to maintain a consistent copy of a centralised style queue at every PE, and have every PE execute the same scheduling algorithm. The techniques used for maintaining consistency in distributed replicated databases could be applied. However, these techniques have high overheads and are probably not appropriate to schedulers. Proposed approaches to decentralised scheduling do not attempt to maintain an absolutely consistent database between all PEs.

Despite the emphasis on load balancing in what follows, it should be borne in mind that a scheduler may be oriented towards not assigning work to failed PEs. Achieving this sort of fault tolerance requires most of the same supporting mechanisms as achieving load balancing.

2.1 Dipstick Methods

The dipstick proposal for load balancing is a conceptually simple method based on the analogy of the oil level dipstick in a car engine. When the load at a PE (as measured by some appropriate metric) lies between the pre-defined ‘MAX’ and ‘MIN’ values, then the PE processes only its own workload. Should the load at a PE fall beneath the ‘MIN’ level then the (executive of the) PE broadcasts to all other PEs that it is able to accept work from them. Conversely, should the load exceed the ‘MAX’ threshold then the PE will export to those PEs willing to take work, until its load level is down beneath the ‘MAX’ point.

A dipstick mechanism was proposed for the Integrated Computer Network System (Howell, 1972) but was abandoned because of anticipated difficulties in providing remote file access for exported jobs. Although nothing has been published, we have been informed that an experimental dipstick mechanism was tried also on the Xerox PARC POGOS system (Redell, 1977). The POGOS system is a local network of 25 Data General Nova minicomputers. It was found to be impossible to set the values of ‘MIN’ and ‘MAX’ so as to produce stable load balancing behaviour. Either no work moved between PEs or (for other settings of ‘MIN’ and ‘MAX’) the communication system was swamped with work being moved around, resulting in very little actual processing being done.

The deficiencies of the dipstick method are:

1. It offers no guidance as to which work should be moved nor, when a choice exists, where it should be moved to. No account is taken of the resource requirements of the moved work.
2. The timing of moves is arbitrary. Computations are moved at any time instead of at special points in their progress. If a computation were moved between finishing one activity and starting another, then there would be less, often very little, data that would have to be shifted with it. Also staggering the times when PEs can offload work gives more time for updated PE status to be propagated, leading to more stable operation.
3. The use of fixed ‘MAX’ and ‘MIN’ values limits the range of loadings for which the policy is useful.

2.2 Bidding

The bidding method of decentralised scheduling is based on the economic model of perfect competition. Within the economic model selection of the lowest priced goods or services leads to the optimum usage of resources. The basic idea of the bidding scheme is that the (executive of) each PE submits a bid to carry out some processing service. The bidding is then carried out on the PE that submits the lowest bid.

This scheme was proposed for one of the earliest distributed computer systems, DCS (Farber, Feldman, Heinrich, Hopwood, Larson, Loomis and Rose, 1972). The following is a brief description. When a new job or task is to be started, the PE responsible for its initiation broadcasts a ‘request for bids’ message to all PEs. This message contains details of the resource requirement of the new task. Each PE works out a bid, its ‘price’ for carrying out the service. The bid is derived from the PE’s current loadings and resources. These bids are returned, as messages, to the initiators which selects the lowest one. The initiator PE sends a message to the chosen PE to indicate it was the lowest bidder. Bids are not binding however (otherwise each PE would have to reserve all the resources involved in each bid until it had been notified of the outcome). The chosen PE decides if it is going to perform the task or not. If so, it notifies the initiator. The initiator sends off any relevant data, and the chosen PE then executes the task. If the resource situation of the chosen PE has materially altered since it issued the bid then it notifies the initiator that it is no longer able to carry out the task in the DCS scheme the initiator then has to start the whole bidding process again by broadcasting a new ‘request for bids’. An alternative action would be to try the second lowest bidder.

There are several attractive features of bidding schemes:

1. Scheduling is done at a time of ‘minimum context’. That is a decision on where to locate a task is taken only on its initiation or when it requires a completely new service. Thus, the amount of data that has to be shifted around when a job shifts to a new PE is minimal.
2. Account is taken of the resource requirements of
the task to be moved. For example, if a certain task requires a peripheral device, then a PE that does not have that (kind of) peripheral will return a maximum size bid for that task.

3. The bidding scheme can also provide a mechanism for increasing the number of replicated copies of code within a system (though this was not proposed for DCS). Other things being equal, a PE that has the code required by a task already in primary memory will bid lower than a PE that would need to load the code from backing store or another PE. If the latter PE was chosen then this would increase the number of replications of the code for that task. There does not seem to be anything published on actual performance of DCS. Difficulties that may be encountered by an implementor of a bidding scheme could include the following:

1. The scheme has very high communication overheads. In a system of n PEs a single bidding cycle consists of a broadcast, n-1 reply messages and the notify and accept/reject messages. As stated above, the impact of this overhead may be on the processing power of PEs rather than on the communications subnet. For each bid, each processor has to interrupt its current task, examine the resource requirements of the new task and formulate and transmit its bid. The total system overhead is of the 'n squared' type. The overhead on each of the n PEs is directly proportional to n (until the overheads significantly impact the rate of initiation of new tasks).

2. The delay in task or service initiation caused by the bidding scheme is excessive. Even if a task is ultimately run on the same PE as its initiator there is the delay of the broadcast and waiting for all responses (or timing out and limiting the choice to only those PEs that have responded). Theoretically there is no limit to the number of times an initiator has to go through the bidding cycle before successfully getting a PE to accept the task.

3. Because bids are non-binding, the scheme may degrade badly under heavy load conditions. It is precisely then that their resource availability will change most rapidly, leading PEs to renegot on their bids and so cause further overheads, prolonging the period of overload.

4. The final difficulty concerns the mapping of the theoretical economic model onto the reality of computing systems. It is difficult to specify how availability of various resources should be translated into a single one-dimensional bid. The perfect competition theory assumes a smooth trade-off between any pair of resources. But the trade-off between processing time and memory usage (for example) is not smooth and, in any case, normally cannot be made dynamically. We have been informed that the final choice of the DCS implementors was to calculate bids as a function only of the free primary memory of a PE.

2.3 Adaptive Learning

A scheme with minimal communications overhead forms part of a proposal for the Distributed Function Multiple Processor, DFMP (Colon, Glorioso, Kohler and Li, 1976). In this system it is assumed that there is a fixed set of services provided.

The DFMP proposal postulated a system consisting of a number of nodes, with each node consisting of a group of functionally specialized PEs. Each PE is dedicated to performing one kind of service for that node with a special PE performing the decentralised scheduling operations. The proposed load balancing mechanism could be easily adapted for multiple computer systems where each PE can perform any service and the scheduling is carried out by the local executive.

In the DFMP system the scheduler at each node maintains a table of the performance of each node for each kind of service. All requests for new services are channelled through the scheduler which selects which node to send the request to. The receiving node queues the request for ultimate execution of the service. By way of an acknowledgement the receiving node scheduler returns to the originator the number of requests it has queued for that service. The originator scheduler uses that information to update the relevant entry in its table.

The crucial part of the procedure is the method of determining the server node from the information held in the table. A naive method would be to always choose the node whose table entry showed the shortest queue. This would produce unstable behaviour. Those nodes not being chosen would have no way of modifying the table entries, which consequently would become totally meaningless.

For the DFMP system it was proposed that the queue length information be used to assign 'probabilities of service' for each node. The choice was then a random draw from a distribution with the corresponding probabilities. Thus the node with the shortest queues would be chosen more often on average than one with longer queues but feedback would eventually be received from all nodes. Whether it is possible to set probabilities such that the choice of server node is not almost entirely random nor almost totally deterministic is an open question.

Certainly there seems to be a mismatch between the underlying Cybernetic theory of learning and its application in this case. In the Cybernetic theory a good choice is 'learnt' by a series of trials in which a good outcome increases the probability of the choice that leads to it and conversely bad choices are discouraged by decreasing their probability. But in DFMP the optimum choice of node in a distributed system is changing as fast as the trials to determine that optimum.

This method does do scheduling at times of 'minimum context'. It also has very low communications overheads. This, however, is the fundamental problem; the delay in updating information is so large that it is almost bound to lead to feedback instabilities.

3. A COMPLETE MECHANISM FOR LOAD BALANCING

3.1 Description of Underlying System Architecture

This section details the thinking behind the decentralised scheduling method outlined in Casey and Shelness (1977). The starting point for this scheme was a distributed computer system that appears to the user as a single system, compared to the DEMOS system (Dowson, Collins and McBride, 1979). A computation is considered to go through stages, demanding a different service and hence using different resources at each stage. Each service may be executed on a different PE but this has to be transparent to the user.
To facilitate the movement of computations between PEs a way had to be found of describing all the currently active components of the computation. For this task, the domains and capability lists of protection work proved ideal. In order to confine a task to its current environment (the basic aim of protection work) one has to be able to describe its environment, that is, the resources such as memory segments and peripherals it currently can access.

Thus, when a computation wishes to invoke a new service, it was assumed that it was possible to construct a list of capabilities, or descriptors, describing the resources needed for the new service. This list of descriptors is called a clist. In a simple case there is a descriptor for the code, one for the process base, one for the workspace segment and one for any parameters passed to the service. The process base segment is used to store registers and other control information for the computation. If there are only a few simple parameters to a service they may be carried in registers in the process base segment, saving the overhead of managing another segment. A clist describes a domain, the total resources required by a computation while it performs a service.

As well as giving the location of the segment, a segment descriptor gives its size. Descriptors for work space segments (to be created when the service is initiated) contain a size only. Descriptors for peripheral devices give the location only. Figure 2 shows a clist for a possible disk handling domain. When the service is requested the segments may be located in different PEs and the PE identification in all descriptors will not be the same. Before actual execution of the disk handling service can begin, all the segments have to be moved to the same PE (and a workspace segment be created there). Figure 3 depicts a clist for a possible disk handling domain. When a service is completed a clist defining the domain, the clist is transmitted to it. Using the information in the clist, this PE requests from other PEs the segments that it does not already have. When all the segments arrive the PE is able to perform the service.

### 3.2 Gathering Status

The scheduling strategy was developed gradually, each modification being simulated and its effect determined (Casey, 1978). Central to the scheme was the acquisition by each kernel of the status of all PEs. This status information consisted of the load and free memory at each PE. The load was defined as the number of ready to run computations at the PE (some simulation runs used this figure plus the number which were waiting for segments to be loaded from other PEs as the load — there did not seem to be any significant difference in system performance).

It was decided that the cheapest way to collect this information was to add an extra two bytes to every message in the system. A sender would place its own load and free space values in a message just prior to transmission. The receiver would use these values to update its status tables. For certain types of communication subsystem, such as Ethernet, it is possible for all PEs to eavesdrop on all passing messages. An intelligent communications interface unit would be able to pick off the source and status bytes of passing messages and update the status tables. Both forms of status gathering were simulated. Message traffic is highly correlated with changes of a PE's status. It was found however that when only the receiver of a message used the status it contained then it was necessary for a PE to broadcast special status messages should it get into a critical state (such as having no free memory space).

### 3.3 The Scheduling Decision Procedure

The initial strategy adopted was pure minimisation of data transmitted between PEs. The rationale for this was that in many distributed systems the total communications bandwidth is fixed irrespective of the number of...
PEs in the system. Conserving bandwidth would thus enhance extendibility. The PE chosen for a service was the one that already had the most space taken by segments for that service. As partly expected, this policy alone proved insufficient to distribute the computing load among all PEs. The simulation experiments showed that computations clustered in a few PEs. Some trouble was also experienced over memory space when the simulated system was heavily loaded.

While retaining the basic idea of minimising the load on the communications subsystem a number of modifications were made. Each segment size is modified by several multiplicative factors before being compared with the others.

1. Segment sizes are decreased by a factor representing the load at their current PE. The effect of this is to make the choice of a lightly loaded PE more likely.

2. Segment sizes are increased by a factor representing the amount of free space at their PE. This is to encourage migration to PEs with plenty of free space.

3. To avoid the constant movement of shared segments between domains located at different PEs, shared segment sizes are multiplied by the count of domains they form part of. This factor encourages the domains that have common segments to collect at one PE.

4. Often a computation involves multiple calls on the same service. By simply keeping the name of the service last called it can be determined if the service currently being invoked is the same as the last one. The size of segments being passed from the calling domain to the called domain (the process base and parameter segments) was then multiplied by the count of consecutive calls. The effect of this is to cause the code of caller and called services to migrate to the same PE.

A fundamental factor limiting load balancing was that eligibility for choice as the PE to perform a new service was restricted to those PEs already holding a segment required for the new service. A second stage in the procedure introduced whereby the PE chosen as the best location for the new service could nominate another PE instead. This was done if the chosen PE determined its workload to be substantially more than some other PE's load. This second level mechanism was also invoked if the chosen PE determined it did not have enough space for all the segments to be imported or created. In this case another PE with enough space was nominated and the clist passed to it. To stop the case of the clist being passed around indefinitely, a limit to the number of times it could be passed around was imposed. (Before this limit was implemented and when status was being acquired from just incoming messages, there was one simulation run which resulted in clists being passed around continuously).

The whole process of choosing a PE on which to perform a new service can be summarised as follows:

**STAGE 1**

The choice is made between PEs that have one or more segments of the new domain. The PE chosen is the one with the largest aggregate size of segments, with segment sizes weighed by factors representing:

- other uses of the shared segments
- the load at the segments' PE
- the free space at the segments' PE
- previous calls of the same sequence of services.

**STAGE 2**

The choice among all PEs is made by PE chosen at Stage 1. This PE chooses:

1. Any PE that is substantially less busy and has the free space, otherwise.
2. itself unless it does not have enough free space, when any site that does have enough free space is chosen.

### 3.4 Sample Load Balance Figures

More simulations were performed with the above modifications in effect. Although they resulted in a big improvement, the workload was still not evenly distributed.

As mentioned in Section 1, it is difficult to give a quantitative measure of the degree of load balance obtained. One of the sets of figures obtained from the simulation studies testing the load balancing mechanism was the percentage idle time of each processor. Table 1 gives these figures for three simulations of a system of eight PEs. Each PE was assumed to be of a power roughly equivalent to a PDP 11/70 and to have 192 Kbytes of memory. A general time-sharing load was assumed with 'saturation' occurring at a load of 10 users per PE. The main peripherals, i.e., disks and consoles, were assumed to be controlled by separate specialised PEs. Each simulation run covered about 1400 user interactions.

Table 1 shows the load balancing mechanism distributing the load very evenly when the system is heavily loaded. When the system is lightly loaded the distribution is less even. But the aim of load balancing is to avoid having work waiting on one PE while another PE is idle. For a lightly loaded system it is not necessary to spread the load evenly to achieve this.

Another figure of interest is the number of times the Stage 2 calculations results in a change in the chosen PE. For systems with the 'saturation' load or greater, about five per cent on average of the Stage 2 calculations resulted in a change of PE. This figure reached 10 per cent.

<table>
<thead>
<tr>
<th>No of Users</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>112</td>
<td>3.8</td>
<td>2.3</td>
<td>2.5</td>
<td>2.6</td>
<td>6.3</td>
<td>1.6</td>
<td>6.8</td>
<td>3.9</td>
</tr>
<tr>
<td>80</td>
<td>22</td>
<td>20</td>
<td>15</td>
<td>23</td>
<td>8.0</td>
<td>14</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>48</td>
<td>32</td>
<td>44</td>
<td>10</td>
<td>34</td>
<td>51</td>
<td>37</td>
<td>42</td>
<td>25</td>
</tr>
</tbody>
</table>

---

for some simulation runs. For lightly loaded systems the Stage 2 calculation resulted in a change of PE in less than two per cent of all cases.

4. CONCLUSIONS

The simulated system described in the previous section has shown that in all likelihood a multicomputer system employing decentralised scheduling could be built. Obviously a system employing such decentralised scheduling need not employ exactly the same mechanisms as those that have been described. However, any successful design will have to pay particular attention to the following points.

1. The need for each PE to have status information relating to other PEs. There is a tradeoff between the overheads of maintaining up-to-date information and the inefficiencies that result from using less current information. The simulated system has demonstrated however that absolute consistency between each PE's view of the total system is not a requirement for success. The opportunity to review scheduling decisions is required (e.g., the Stage 2 calculation above).

2. The granularity of load balancing. Indisputably, moving work around a system incurs high overheads; the question is, at what frequency of movement do these overheads outweigh the gains of utilising otherwise idle processors?

3. The total resources required by each unit that can be moved around: A scheme that takes no account of the behaviour of the computations it moves is inviting inefficiency in resource usage.

Newer designs for multicomputers are tending to place a lot of intelligence into communications interfaces. This processing power could be used to gather status information and to assist in the movement of data between PEs. Decentralised scheduling could become a reality on some of these systems.

REFERENCES


For author's biography see page 145 of the November 1980 issue.
HYPERDISK, AN ACCESS METHOD FOR REMOTE DISK DEVICES

George Havas*

*The author is with CSIRO Division of Computing Research, Canberra. Manuscript received 20 March 1981.

A method currently under development for accessing remote IBM-compatible disk devices from IBM-compatible processors is described. The method is transparent to the application programmer and utility user. It provides for data transfer over the Network Systems Corporation HYPERchannel which is the basis of the CSIRONET local computer network.

Keywords and Phrases: remote disks, local computer networks, resource sharing.


1. INTRODUCTION

The HYPERdisk project is being undertaken with the co-operation of CSIRO, FACOM Australia Limited, Techway and Storage Technology of Australia. The aim of the project is the development of a method for accessing remote IBM-compatible disk devices from IBM-compatible processors running IBM MVS or FACOM OSIV/F4 operating systems. The initial implementation work is being done on the CSIRONET local network, a description of which is given in Wolfendale (1980), where this project is briefly mentioned (section 5) and further references may be found.

The access to remote disks utilises an Assembler program, referred to as HYPERD, which runs under an IBM MVS or FACOM OSIV/F4 operating system. A minimal hardware configuration for running HYPERD includes an IBM-compatible processor connected via a Network Systems Corporation (NSC) A220 adapter to a HYPERchannel together with an IBM-compatible disk controller and disks connected to the HYPERchannel via an NSC A510 adapter. The HYPERD software is designed to run on multiple hosts attached to a HYPERchannel network so that any host can access any disk which has a controller connected to the network.

Since the initial implementation is being done using a FACOM M190 host processor this description is slanted in terms of the FACOM OSIV/F4 operating system. The primary difference in HYPERD software for IBM MVS systems is due to the fact that IBM channel programs use only real addresses whereas FACOM channel programs allow virtual addresses. This affects some technical details in the analysis of the syntax and semantics of channel programs but does not affect any of the underlying principles.

2. SOFTWARE OUTLINE

The HYPERD software relies on the fact that the FACOM OSIV/F4 (and the IBM MVS) operating system has a Basic I/O Supervisor (BIOS) through which virtually all input/output is done. A feature of BIOS is that it includes only a very small number of machine I/O instructions. This feature is utilised by HYPERD which provides for the interception of all the BIOS I/O instructions.

It is worth noting that NSC has already developed software for accessing non-disk peripherals on the same kind of network (see Hardwick [1980]). The primary difference in philosophy between the approach adopted by NSC and the approach described here is that the NSC non-disk software relies upon I/O interception at a higher level. This necessitates some source level changes to the operating system, whereas the lower level interception method used by HYPERD requires no changes to the operating system beyond HYPERD itself.

HYPERD carries out the following functions.

2.1 HYPERD establishes the primary interception system, initialising the nucleus level code and the run-time data base.

2.2 HYPERD initialises the remote disk tables and tests communications via the HYPERChannel to ensure integrity of the system.

2.3 HYPERD activates primary interception, whereupon remote I/O may commence. The remote devices are established as online to the host processor.

2.4 Once interception is activated all BIOS I/O instructions are trapped. I/O instructions for all devices other than remote disks are performed normally with only minimal overhead. I/O instructions for the remote disks are not executed directly, but rather a secondary task is scheduled which processes the I/O over the HYPERchannel.

2.5 For start-I/O instructions the secondary task analyses the syntax and semantics of the associated channel program. It constructs corresponding channel programs for communication between the host processor and the A220 adapter and between the A510 adapter and the disk controller, and supervises the execution of these new channel programs.

2.6 HYPERD can be stopped at any time, whereafter the remote devices are offline to the host processor.

3. FUNCTIONAL AND PERFORMANCE LIMITATIONS

In principle HYPERD can handle all disk I/O which goes through BIOS. However the NSC A510 adapter has various restrictions on the types of channel programs that it can handle. In particular it cannot handle self-modifying channel programs or channel programs with more than 63 channel command words. Disk channel programs which violate these restrictions will not be executed via HYPERD, but an error will be signalled. There should be minimal impact from these restrictions.

As far as MVS and OSIV/F4 operating system access methods are concerned there is only one method not supported by HYPERD, namely ISAM. This is because ISAM issues self-modifying channel programs. However ISAM is obsolete, being replaced by VSAM, so inability to support ISAM is not a substantial drawback. All other current access methods are supported.

The current A510 design is such that certain types of channel programs will have their efficiency impacted. For example, the A510 does not support PCI interrupts in the same way as a normal channel. Channel programs for remote disks which use PCI interrupts will have the interrupt presented late. In the context of the MVS and
OSIV/F4 operating systems this does not lead to erroneous results but may reduce efficiency. Also the reading (writing) of adjacent records on a disk via the A510 leads to a lost disk revolution, a situation which will be handled automatically by the disk controller, but with efficiency downgraded.

A census of the day to day traffic on the CSIRONET FACOM M190 reveals that over 80 per cent of disk I/O will not be troubled by inefficiencies due to the current A510 design. Subsequent stages of this project will be directed to software and hardware redesigns for the A510 which will eliminate or minimise efficiency loss.

Running HYPERD on multiple hosts on the same network gives each host access to the networked disks. However HYPERD does not support a management system for orderly access. Multiple access is to be controlled by operational procedures and host-dependent software.

REFERENCES
Letters to the Editor

FURTHER DISCUSSION ON JOURNAL CONTENTS

I write in support of the Editor's choice of articles. Over the several years in which I have been an ACS member, I have read each issue of the Journal, and found the articles to range from interesting and helpful to completely obscure; but that presumably reflects my interest and knowledge, not the ability of the authors. In reading the letter of Mr McGuinness in the February issue, I find particularly irritating his repeated use of the phrase, "the majority of... in industry". As things stand now the Society attracts very few members outside of the business community. If we want participation by all types of users, then articles cannot be limited to business uses of COBOL or similar material of interest to a subset of the industry. Further, the particular article chosen by Mr McGuinness for his attack is the only one I have seen in the Journal that I have saved and called to the attention of other people.

J.M. Graham,
Communication & Electronic Engineering,
Royal Melbourne Institute of Technology

EDITOR'S COMMENTS

I wish to thank Mr Graham for his support of the Journal's editorial policy, but would like to make one or two clarifications. First, normally papers are accepted for publication only after assessment by referees with expert knowledge of the subject. Thus, the contents of the Journal reflect not so much the personal judgement of the Editor; rather, the supply of material and the views of the referees. And there is a simple reason why most of the material we receive is on specialized research topics: People engaged in research have a constant incentive to publish because new research findings are of no value until they have been widely disseminated, their validity independently verified and their utility proven in experiments and applications elsewhere. This also explains why so much of the material is rather preliminary and not directly related to the daily concerns of practitioners. Take, for example, the early papers of E.F. Codd on relational databases, with their abstract terms, notations and theorems, and having apparently no relation with database processing as it was then practised. What a loss it would have been if IBM, his employer, had refused to support such "esoteric" stuff! Without claiming that the papers published in the Journal are quite as important as those of Codd, I would nevertheless insist that their publication is a worthy cause, and ACS members should not grudge the few dollars per year it costs them. Perhaps many members would prefer to have a publication like Datamation, but that would not be possible without full-time staff writers and paid authors, and financial outlays of a type which, I suspect, few members are prepared to support. However, I would encourage anyone who disagrees with the present policies to make his voice heard, and would welcome any concrete proposals about how to improve the Journal.

We should like to endorse some of the comments made by Mr G.E. McGuinness re the nature of articles published in the Australian Computer Journal. We believe we belong to the majority group of members who on most occasions experience difficulty in understanding both the TITLE and CONTENT of the published articles. We suggest that the publication be circulated to the minority group who see value in it. For those members who do not wish to receive a copy we suggest their fees be rebated by the $2 quoted in your letter.

R.W. McDonald, MACS
P.M. Waters, AACS
R.J. Cooke, AACS
Software Holdings Ltd,
Mount Gambier, SA 5290

EDITOR’S COMMENTS

ACS members who share this view are advised to approach their Branch Executives and the National Council, which alone decides whether the suggested scheme ought to be implemented. I welcome the fact that members who disagree with present policies are speaking out and suggesting concrete changes.

COMMENTS ON "NEBALL AND FINGRP: NEW PROGRAMS FOR MULTIPLE NEAREST – NEIGHBOUR ANALYSES"

This paper includes the result that the random expectation of Or, defined by \( \frac{1}{n} \times \{ \text{the number of elements which are members of the first r neighbours of both of two fixed elements j and k} \} \) where n is the number of elements, is approximately:

\[
e_r = \frac{1}{n} \sum_{s=1}^{r} \left( \frac{r}{n} \right) \left( \frac{n-1}{n} \right)^{r-s}. \tag{1}
\]

This equality is exact under the assumptions of random ordering of neighbours and independence between lists of neighbours for each element. Abel and Williams missed the simplification that under the assumptions,

\[
e_r = \left( \frac{r}{n} \right)^2. \tag{2}
\]

A simple proof of this is;

\[
O_r = \frac{1}{n} \sum_{i=1}^{n} X_i \text{ where } X_i = \begin{cases} 1 & \text{if the i-th nearest neighbour of k is among the first r neighbours of j} \\ 0 & \text{otherwise} \end{cases}
\]

\[
E \left[ X_i \right] = \frac{r}{n} \text{ and } E \left[ O_r \right] = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{r}{n} \right) = \left( \frac{r}{n} \right)^2.
\]

Hence, \( e_r = E \left[ O_r \right] \).

This simplification considerably quickens the computation of \( e_r \). As pointed out in the original paper the above expres-
Letters to the Editor

The exact expression for $e_r$ is slightly biased since the nearest neighbour of element $j$ is constrained to be $j$. Under this restriction,

$$E[X_1] = \sum_{i=1}^{r} p \left[ \frac{\text{element } k \text{ is among the first } r \text{ neighbours of } i}{\text{element } j} \right]$$

For $i = 2, \ldots, r$,

$$E[X_i] = \sum \left\{ \begin{array}{c} \{ \text{ith nearest neighbour of } k \text{ is among the first } r \text{ neighbours of } j \} \\ \{ \text{ith nearest neighbour of } k \text{ is among the first } r \text{ neighbours of } j \} \end{array} \right\}$$

$$= \sum \left\{ \begin{array}{c} p \left[ \frac{\text{ith nearest neighbour of } k \text{ is among the first } r \text{ neighbours of } j}{} \right] \\ \{ \text{ith nearest neighbour of } k \text{ is not among the first } r \text{ neighbours of } j \} \end{array} \right\}$$

$$= \sum \left\{ \begin{array}{c} \{ \frac{(r-1) x \frac{r-1}{n-1}}{n-1} \} + \{ \frac{r}{n-1} x \frac{n-r}{n-1} \} \\ \{ \frac{(r-1)^2 + r(n-r)}{(n-1)^2} \} \end{array} \right\}$$

Hence,

$$e_r = E[O_r] = \frac{1}{n} \sum_{i=1}^{r} E[X_i]$$

$$= \frac{(r-1)(n-2r+n)}{n(n-1)^2}$$

$$= \frac{(r^2 - (1 - \frac{r}{n})^2)}{(n-1)^2}.$$  (3)

There are plans to incorporate the "exact" expression in the NEBALL program (Abel, personal communication).

The maximum possible value of $O_r$ for fixed $r$ is $\frac{n}{2}$ and the minimum possible value is:

$$\min O_r = \begin{cases} 0 & \text{if } r < \frac{n}{2} \\ \frac{2r}{n} & \text{if } r \geq \frac{n}{2} \end{cases}.$$  

since for $r \geq \frac{n}{2}$ the configuration which gives $\min O_r$ occurs when the $n-r$ furthest neighbours of $j$ are completely disjoint from the $n-r$ furthest neighbours of $k$. Using these results and (2) it can be shown analytically that the approximate formulation of $\Delta_r (=O_r-e_r)$ is constrained between $\pm \frac{1}{4}$. Similarly, using these results and equation (3) it follows that the "exact" formulation of $\Delta_r$ is bounded above by $\frac{1}{4}\left\{ 1 + \frac{(n-1)^2}{(n-1)^2} \right\}$ if $n$ is even and $\frac{1}{4}\left\{ 1 + \frac{(n-1)^2 + 1}{(n-1)^2} \right\}$ if $n$ is odd.

From equation (3) it can be seen that the bias of the approximate $e_r$ given in the Abel and Williams paper is $O\left( \frac{1}{n^2} \right)$. For $n=25$, the bias for various $r$ is:

<table>
<thead>
<tr>
<th>$r$</th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>bias</td>
<td>.0016</td>
<td>.0011</td>
<td>.0006</td>
<td>.0003</td>
<td>.0001</td>
<td>0 exactly</td>
</tr>
</tbody>
</table>

It is clear that although the bias is small, for moderate $r$, it is greater than "of the order $10^{-4}$" as stated in the Abel and Williams paper.

Finally, I would like to query their statement that they are using a Kolmogorov-Smirnov type statistic. A one-sided Kolmogorov-Smirnov statistic (Siegel, 1956, p. 48) compares an observed cumulative distribution with a theoretical distribution and not with a set of expected values, and a two-sided Kolmogorov-Smirnov statistic (Siegel, 1956, p. 128) compares two independent observed cumulative distributions.

References


K. L. Butler, CSIRO Division of Mathematics and Statistics, Cunningham Laboratory, St. Lucia, Qld. 4067

AUTHOR'S REPLY

We congratulate Mr Butler on his analysis of the problem. The exact expression for $e_r$ has been incorporated in NEBALL.

D.J. Abel, CSIRO Division of Computing Research, Townsville, Qld. 4810

CALL FOR PAPERS

SPECIAL ISSUE ON DATABASE MANAGEMENT

The Australian Computer Journal will publish a special issue on “Database Management” in November 1981. The issue will be under the Guest Editorship of Professor A.Y. Montgomery, Department of Computer Science, Monash University.

Research papers, tutorial articles and industry case studies on all aspects of computer databases and information systems are welcome, either as full papers or as short communications. Prospective authors should write to:

Professor A. Y. Montgomery, Guest Editor
ACJ Special issue on Database Management
Dept. of Computer Science,
Monash University,
Clayton, Victoria 3168
Australia

to notify him of their intention to submit material for the issue and provide a brief summary of their intended contribution. To allow adequate time for processing, full papers must be received by 1 August 1981, and short communications by 15 August 1981. Material received after the deadlines may be considered for publication in later issues of the Journal.

Papers should be prepared in accordance with the guidelines published in the May 1980 issue of the Journal.

CALL FOR PAPERS

SPECIAL ISSUE ON SOFTWARE ENGINEERING

The Australian Computer Journal will publish a special issue on “Software Engineering” in early 1982. Research papers, tutorial articles and industry case studies on all aspects of the subject will be welcome, and both full papers and short communications will be considered.

Prospective authors should write to:
Professor P.C. Poole, Guest Editor,
ACJ Special Issue on Software Engineering,
Department of Computer Science,
University of Melbourne,
Parkville, Victoria 3052,
Australia.

to notify him of their intention to submit material for the issue and provide a brief summary of their intended contribution.
3M AUSTRALIA PTY LTD REVOLUTIONISES ELECTRONIC FILING

Six second filing and retrieval in less than 30 seconds are the abilities of 3M Australia Pty Ltd's Micrapoint I Electronic Filing System.

The Micrapoint, a self-contained electronic filing system, is an office machine which employs microprocessor technology and is simple to operate, requiring no special skills.

Anyone from typist to office manager can use the system which microfilms the information and stores indexed addresses on magnetic diskettes (floppy disks).

The system comprises a microprocessor with keyboard and video console and uses two diskettes to store index information.

A 3M Page Search Printer and a 3M PST Systems interface may be linked to the Micrapoint.

After a document is filmed, each microfilm frame is indexed from the keyboard with letters or numbers depending on the user's needs.

Retrieval is a simple matter of pushing the appropriate keys; film cartridge and frame numbers for specific documents are then instantly displayed on the video screen, and the 3M Reader/Printer can be directed to automatically display the frames in sequence, or as selected by the operator.

The system's double-density diskettes can store index information for up to 85,000 documents.

Diskettes can be changed in seconds and the filing capacity expanded by the insertion of two fresh diskettes.

DEFENCE COMPUTING INFRASTRUCTURE STUDY

The Minister for Defence, Mr D.J. Killen, said recently his Department had initiated a major study of the future computing needs of the Defence Force. The study began in February 1981 and was expected to last for six months.

Mr Killen said the Department of Defence had been a national leader in computing its administration in the early 1960s. Currently the major pay, personnel, logistics and technical services functions within the Department were heavily dependent on computers.

Over the last five years these systems had been progressively transferred from computers purchased in the early 1960s to newer equipment. This work was planned to be completed in the 1982/83 financial year.

Mr Killen said that over the last ten years there had been a revolution in computing technologies. In relative terms the cost of computers had decreased dramatically.

Today there were few Defence activities which could not benefit substantially from computer support.

The challenge facing any large organisation such as Defence was how best to take advantage of the new possibilities without creating new difficulties.

A wide proliferation of computers could, if not properly managed, lead to later problems in the efficient exchange of information between computers and in the compilation of data for senior management.

The review, known as the Computing Infrastructure Study, was being undertaken by senior professional staff of the Department's Computing Services Division and senior members of the Armed Services. A senior officer from the Public Service Board had also been seconded to the study.

NEW COMPUTER STORAGE FURNITURE

A new range of computer furniture has been launched by the office furniture division of Namco Industries (Vic.), covering storage units for output and input files, and tables for the hardware with accompanying chairs.

There are a variety of specially designed units to store all those printouts - filing cabinets, trolley files, 'Computastore' cupboards, 'Computashelf' shelving, and desks with printout filing drawers - all in a choice of 14 colours plus mix and match two tones.

In addition there is a range of Operators Tables on which to stand machines, with chairs also for the operators, together with a number of different units for the storage of discs, tapes, cards, etc.

Full details are available on request from Namco Office Furniture Victoria.

COMPUTER SYSTEMS TO PEOPLE'S REPUBLIC OF CHINA

An agreement under which the People's Republic of China will buy, up to $5,000,000 worth of Z80-based small business microcomputers from Zilog Inc, has been signed. Initial shipments have already begun.

According to Martin Cohen, Zilog's Vice President of field operations, the Beijing Automation Technology Research Institute has agreed to purchase Zilog's MCZ 1/50 Microcomputer Systems and microcomputer board-level products.

The Institute intends to use the Zilogs systems in airline and hotel reservations systems, as well as for educational purposes at universities. Zilog is providing training for several People's Republic representatives in using the MCZ systems and associated software.

The MCZ 1/50 is a general-purpose desk-top microcomputer system based on Zilog's Z80 microprocessor; it includes a visual display unit terminal and floppy disk storage.

Zilog Inc, an affiliate of Exxon Enterprises, manufactures and markets microcomputer circuits and memory devices, boards, complete systems and associated software.

For further information, contact the Australian Distributors, ZAP Systems Pty Ltd, St Leonards, NSW.

WORD PROCESSOR TRAINING MADE EASY BY PHILIPS

With the first order for 14 Philips P5002 word processing systems and peripherals valued in excess of $200,000 and now installed with the Brisbane City Council, operator training as a second phase has proceeded smoothly with the use of an audio-print training technique.

While individual operator training on a Philips system normally takes five days, 28 operators have been able to be totally trained in using the P5002 system in a period of only one month with a special audio-print training technique developed by Philips. This includes optional programme modules that the system handles including records management, special sorting routines and mathematical calculations. Examples of applications for these programme modules are illustrated by "Personnel Files" management; sorting routines for stock files in supplier groups, age of stocks and pricing categories; and automated mathematical calculations for Monthly Profit and Loss Statements, budget comparisons or complicated equations leading to tables of figures.
EMERY COMPUTERLINK GOES ON DISPLAY

Emery Air Freight Corporation put its "Emcon" computer controlled package tracking system on display at the Software Showcase in Sydney during February.

A full-speed telex link to Emcon from the Emery booth at the Sydney Town Hall exhibition, showed visitors how the $560 million-a-year air freight forwarding organisation maintains its world-wide reputation for speed and efficient service.

Through telex links at Emery offices around the world, the Emcon nerve-centre in Wilton, Connecticut, USA, can track any shipment in the corporation's network and give a status report on that shipment.

Emery NSW regional manager, Mr John Burford, said at the software showcase, that it was far from unusual for the company to go on display to the computer industry.

"More than one in six of the companies exhibiting at the exhibition are regular Emery customers, and we are keen to show the industry how we can help them," he said.

"We can ship software on paper or stored on various media, virtually anywhere in the western world -- and from the USA to China.

"The speed and efficiency we offer are the factors which attract the computer industry to Emery. We also offer a customs brokerage service.

"That attraction is not restricted to software shippers. We have handled a good deal of hardware into Australia over the years," he said.

The Emery display used at Software Showcase was shipped from the USA and featured printed and video taped information about the company, as well as the telex link.

COMPUTERACC TO HANDLE MINI GAMES RESULTS

For the past four months Brisbane-based software group ComputerAcc have been working under contract to design and co-ordinate a suitable results network for this year's October Mini Games.

The first details of their proposed network were revealed at a Press Club luncheon in Brisbane recently, following the launch of the official Commonwealth Games mascot.

Based on existing programmes used at the 1978 Edmonton Games, ComputerAcc's software specialists have begun the task of systems modification and data up-dating required to produce a compatible and comprehensive results dissemination package.

The 'results network' as it will be known will take in Brisbane's eleven separate games venues, relaying confirmed details for each sporting event to ComputerAcc's central on-line bureau facility at Spring Hill.

Handling the task of storage, collation and dissemination will be a Digital PDP 11/70 with 512K bytes main memory, one of four 11/70's that ComputerAcc use in their on-line bureau service.

Major functions of the system will include relay of completed event score data to venues for interfacing with the host broadcaster's coverage; interfacing of confirmed results with venue and central scoreboard systems; detailing of event schedules with updated programme changes; and transfer of data to the media information centre located at Woolloongabba.

Around 20 input terminals will be located throughout the network with printers, allowing instant access to results data by officials, teams, sporting associations and media.

For the media this will prove a particularly useful aspect of the system, allowing instant recall of comprehensive details on athletes, countries, previous records, etc. A demonstration of this instant access feature was given at the luncheon with emphasis on the ease of data recall.

SMALL BUSINESS COMPUTER UTILITY OPENED

A new computer utility for the small businessman was launched in Sydney recently by the Minister for Business and Consumer Affairs Mr John Moore.

ODP On-Line Data Processing is unique in Australia because it is the only computer utility which caters specifically for the small and medium businessman. Mr John McEnearney, managing director of On-Line Data Processing said, "ODP's Computer packages mean that for the first time businesses as small as the local newsagent through to the medium sized manufacturer can take advantage of hundreds of thousands of dollars worth of usable technology, at a price which remains cost-efficient with the size of the operation."

"By operating a terminal out of the businessman's own office or from a facility such as at the Sydney Chamber of Commerce, the small to medium businessman can use the computer facility to speedily update and process information which may otherwise take days to correlate," he said. "It is common knowledge that computers in big business have saved precious time and money.

"Now this can be the prerogative of the small-medium businessman.

"The ODP Utility will save him time, give him better viability, leading to greater profits," Mr McEnearney said.

"The software programmed by ODP into the computer Utility is regarded as being among the best in the world.

"It is far more powerful, time-saving and efficient than many other more expensive pieces of computer technology," he said.

ODP is designed to play a most important role in the development of Australia's largest commercial enterprise -- Australian small business.

FLAT TUBE IS KEY TO "HANGING PICTURE" TV

The long awaited flat tube television which will eventually hang on the wall like a picture is to go into production next year.

Sinclair Research Ltd (SRL), of Cambridge in eastern England, has announced the successful completion of a five-year research and development programme on a flat TV tube that is just 19 millimetres deep.

First use of the revolutionary tube will be in a new pocket TV and FM radio that will sell for around $100 when it is introduced next year. The Microvision set will be produced in a highly automated plant at the Dundee, Scotland.

The first phase of a $10 million four-year capital investment programme will cover the establishment of an advanced manufacturing facility at Dundee that will employ 250 people next year and be able to produce a million flat tubes a year. By 1985, 1000 people will be working at an expanded plant that will have a capacity to make several million tubes.
Because he's so touchy, you have to treat your computer with kid gloves to get the best out of him.

At Cemac Interiors we do everything from keeping him happy to spoiling him rotten.

**Static gives him the hiccups:**
Unsuitable floor coverings cause static charges that go to earth via a computer. And when he hiccups you get data loss or complete shut-down. To prevent this we install Cemac Tate anti-static carpet as part of our Cemac Tate Access Flooring System.

**Access Flooring not only takes care of unsightly wires, cables and plumbing, but saves considerably on re-location time and expense. You have total flexibility with telephones, power outlets, lights, gas and water pipes and will save on central air conditioning costs.**

**Dust makes him sneeze:**
Plaster and mineral fibre ceilings give off dust that gets into your computer and onto micro-chips, diodes, transistors and electrical connections. And when he sneezes it causes data loss and interruption, short circuits and machine shutdowns. So above their access floors Cemac Interiors install metal pan ceilings that don't make him sneeze.

**Heat makes him blow his cool:**
If a computer gets too hot it shuts itself off. So when he blows his cool, the result is very expensive downtime. That's why Cemac Interiors also offer single source responsibility for the installation of air conditioning and fire protection equipment as part of a total access flooring package.

**Rip-offs leave him completely devastated:**
If the confidential data he's racked his brain to produce falls into the wrong hands, he gets very upset. So Cemac Interiors arrange total security protection.

Cemac Interiors also supply ambient and task lighting, partitioning, and computer room work stations from the Westinghouse Open Plan Office System.

In fact everything to keep your computer working happily.

Give us a call and we'll put it all together for you. **Beautifully.**
The Australian Computer Journal is an official publication of the Australian Computer Society Incorporated.


EDITORIAL COMMITTEE: Editor: C.K. Yuen, CSIRO Division of Computing Research, PO Box 1800, Canberra, ACT, 2601. Associate Editors: J.M. Bennett, T. Pearcy, P.C. Poole, A.Y. Montgomery, J. Lions.

SUBSCRIPTIONS: The annual subscription is $18.00. All subscriptions to the Journal are payable in advance and should be sent to the Australian Computer Society Inc., PO Box N26, Grosvenor Street, Sydney, 2000. A subscription form may be found below.

PRICE TO NON-MEMBERS: There are now four issues per annum. The price of individual copies of back issues still available is $2.00. Some already out of print. Issues for the current year are available at $5.00 per copy. All of these may be obtained from the National Secretariat, PO Box N26, Grosvenor Street, Sydney, 2000. No trade discounts are given, and agents should recover their own handling charges.

MEMBERSHIP: Membership of the Society is via a Branch. Branches are autonomous in local matters, and may charge different membership subscriptions. Information may be obtained from the following Branch Honorary Secretaries. Canberra: PO Box 446, Canberra City, ACT, 2601. NSW: Science House, 35-43 Clarence St, Sydney, NSW, 2000. Qld: Box 1484, GPO, Brisbane, Qld, 4001. SA: Box 2423, GPO, Adelaide, SA, 5001. WA: Box F320, GPO, Perth, WA, 6001. Vic: PO Box 98, East Melbourne, Vic, 3002. Tas: PO Box 216, Sandy Bay, Tas, 7005.

CONTRIBUTIONS: All material for publication should be sent to the Editor for processing. Prospective authors may wish to consult manuscript preparation guidelines published in the May 1980 issue. The paragraphs below briefly summarise the essential details.

Types of Material: Four regular categories of material are published: Papers, Short Communications, Letters to the Editor and Book Reviews. Generally speaking, a paper will discuss significant new results of computing research and development, or provide a comprehensive summary of existing computing knowledge with the aim of broadening the outlook of Journal readers, or describe important computing experience or insight. Short Communications are concise discussions of computing research or application. A letter to the Editor will briefly comment on material previously appearing in the Journal or discuss a computing topic of current interest. Descriptions of new software packages are also published to facilitate free distribution.

Refereeing: Papers and Short Communications are accepted if recommended by anonymous referees. Letters are published at the discretion of the Editor, and Book Reviews are written at the Editor's invitation upon receipt of review copies of published books. All accepted contributions may be subject to minor modifications to ensure uniformity of style. Referees may suggest major revisions to be performed by the author.

Proofs and Reprints: Page proofs of Papers and Short Communications are sent to the authors for correction prior to publication. Fifty copies of reprints will be supplied to authors without charge. Reprints of individual papers may be purchased from the Printer (Publicity Press). Microfilm reprints are available from University Microfilms International, Ann Arbor/London.

Format: Papers, Short Communications and Book Reviews should be typed in double spacing on A4 size paper, with 2.5cm margins on all four sides. The original, plus one copy (preferably two for Papers) should be submitted. References should be cited in standard Journal form, and generally diagrams should be ink-drawn on tracing paper or board with stencil or Letraset lettering. Papers and Short Communications should have brief Abstracts, Keyword lists and CR categories on the leading page, with authors' affiliations as a footnote. The authors of an accepted Paper will be asked to supply a brief biographical note for publication at the end.

This Journal is Abstracted or Reviewed by the following services:

Publisher Service
ACM Bibliography and Subject Index of Current Computing Literature.
ACM Computing Reviews.
AMS Mathematical Reviews.
ENGINEERING INDEX INC. Engineering Index.
INSPEC Computer and Control Abstracts.
INSPEC Electrical and Electronic Abstracts.
SPRINGER-VERLAG Zentralblatt fur Mathematik und ihre Grenzgebiete.

Copyright © 1981. Australian Computer Society Inc.

Production Management: Associated Business Publications, 28, Chippen Street, Chippendale, NSW, 2008. Tel: 699-5601, 699-1154. All advertising enquiries should be referred to the above address.