GEOGRAPHIC INFORMATION SYSTEMS PERFORMANCE IN A BROADBAND COMMUNICATIONS ENVIRONMENT

by

David J. Coleman, B.Sc.Eng, M.Sc.Eng.

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
of
Doctor of Philosophy
in the Department
of
Surveying and Spatial Information Sciences

THE UNIVERSITY OF TASMANIA
February, 1994

DECLARATION

Except as stated herein, this thesis does not contain any material which has been accepted for the award of any other degree or diploma in any university nor, to the best of my knowledge and belief, does it contain any copy or paraphrase of material previously published or written by another person, except where due reference is made in the text of this thesis.

signed: David J. Coleman

ABSTRACT

Developments now taking place in computer hardware and broadband data communications promise to provide users hundreds of miles apart with access to the same equipment and data at comparable levels of performance. However, before adopting such technology, both suppliers and potential customers in the spatial data handling community require a clear understanding of potential network usage and the performance, capacity and cost tradeoffs involved.

This research tests the hypothesis that broadband communication networks possess the performance necessary for organisations to manage their geographic information system (GIS) software and databases from a single location while maintaining satisfactory response times to end-users. Approaches to determining representative GIS operations and network usage patterns within an organisation are proposed and tested under actual operating conditions. Controlled experiments measuring GIS performance across high-speed metropolitan area networks are then described, and the corresponding results are compared for different organisational configurations, over varying distances and under different simulated loading conditions.

Experimental results using Telecom Australia's 10 Mbit/sec *FASTPAC* service indicate that satisfactory file transfer and GIS performance can usually be maintained even over long distances. Specific Network File System (NFS) characteristics which adversely affect response-time performance under certain conditions are identified, and operational tradeoffs are assessed using a model of the current *FASTPAC* tariff structure. Finally, the implications and applications of broadband communications networks in the spatial data handling community are discussed and possible extensions to the research are identified.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This thesis is the result of the efforts, guidance and patience of many persons over the past three years. While it would be impossible to acknowledge all the individuals and groups who have supported this research, I would like to extend my special thanks to the following individuals:

* to Peter Zwart, my supervisor, who patiently kept me on track and reminded me by his example that — while good organisations usually require a combination of good "inside people" and "outside people" — there are still a few individuals that can successfully fill both roles;

* to Michael Montgomery and Andrew Gunter of the Fast Packet Services Group at Telecom Australia in Melbourne, for their continuing financial and moral support throughout this research;

* to Bert Borg at Telecom Australia's Research Laboratories in Clayton, Victoria, whose quiet and competent support provided the magic to make it happen;

* to Simon Greener, Rob Driessen, Julian Dermoudy, Graeme Greenwood at the Centre for Spatial Information Studies — as well as Peter Newton of CSIRO and Marina Cavill of Telecom Australia — whose ongoing advice helped shape the direction of this research;

* to Jennifer and Christopher, whose very real sacrifice of leaving friends behind was hopefully offset by making new friends and, ultimately, by learning how much they could depend on their family and themselves;

* and finally to Betsy, my partner and best friend, whose own sacrifices and unfailing support throughout this adventure are only now truly appreciated, and whose talents and efforts at keeping the family on track paid off in so many ways. I hope I can return the favour someday.

Finally, these investigations were made possible through the financial support and expertise of the Fast Packet Services Group of Telecom Australia in Melbourne. This support – as well as the ongoing assistance provided by the GIS Group at Telecom Research Laboratories in Clayton, Victoria – is gratefully appreciated.

1.0 INTRODUCTION

Developments now taking place in computer hardware technology and broadband data communications promise to provide users hundreds of miles apart with access to the same equipment and data at comparable levels of performance. This has important ramifications to both individual users, who may only wish to use such networks to draw large data files from a remote location, and large programme-driven organisations wishing to maintain centralised responsibility for data and system administration over a growing number of users in different locations.

The growing use of *local* area networks to connect people, computer resources and information within a single location has been an important development in end-user computing over the past ten years. To date, however, very few organisations use *wide* area communications services to routinely link geographic informations systems (GIS) users in remote sites together. The inherent limitations and high costs involved — together with other operational and institutional factors — have compelled most organisations to consider other approaches to GIS data delivery, data management and end-user computing at remote sites.

The goal of this research is to determine whether broadband communication networks possess the performance necessary to enable organisations to satisfactorily manage their spatial information management software and databases from a single location. (The actual hypothesis and its qualifiers will be discussed in detail in Section 1.2.) Through the extension of existing information broadband telecommunication networks in different organisational configurations, over varying distances and under simulated loading conditions.

1.1 CONTEXT OF THE RESEARCH

1.1.1 From Host-Based Systems to Network Computing

The hardware and software originally required for database management systems in general and geographic information systems (GIS) in particular were both complex and insular. Early host-based efforts like the Canadian Geographic Information System [Tomlinson, 1967] operated on larger mainframe or mini-computers and required well-trained systems staff to optimize procedures, develop custom enhancements and manage the data contained within. This environment continued to dominate GIS installations through the 1970's and 80's in forestry, utilities, municipalities and land records management organisations.

In most cases, such host-based systems kept the onus for system and data administration off the end-users and in the hands of experienced data processing specialists. Centralisation of the database also meant greater control over data integrity, since only a single copy of the database was maintained on-line with appropriate security and backup procedures. However, performance of such systems would often degrade in unpredictable ways when more and more users demanded system resources and access to the database. As well, conflicts with system administrators over development and maintenance priorities often resulted in dissatisfaction among end-users in many large organisations.

By 1986, PC-based GIS packages had begun moving geoprocessing out of the hands of information system managers. Besides their low cost, these systems offered more predictable response-times since the user was the only one on the system. Two important disadvantages to GIS users were also encountered, however. First, it was much more difficult to share data among several different people in the organisation. As well, the PC user often had to become his own system and database administrator [Miller, 1990]. While PC-based systems undoubtedly accounted for the dramatic growth in GIS usage through the late 1980's [Zwart et al, 1991], they also put greater onus on managers in large organisations to effectively keep track of the data being collected and processed by an increasing number of end-users with little experience in routine data management procedures.

By the late 1980's, higher-performance workstations connected through local area networks (LANs) became a viable alternative to earlier host-based and stand-alone systems. In this arrangement, GIS users could access and share data as required from a central server where it could be managed effectively. Because computing was distributed to individual workstations, the users could still be assured of satisfactory response times in their own GIS operations [Miller, 1990].

1.1.2 Data Access And Delivery Concerns

The data management problems inherent in LAN-based work groups may be manageable when all system users reside in a single location. However, data communication issues often become problematic when the corporate computing environment grows to include many widely-scattered regional offices. Three important communications issues are often faced by system managers as GIS efforts mature and the nature of system usage changes:

(1) How best to ensure controlled on-line access to all major corporate information systems and archives to appropriate users throughout the organisation?

(2) How best to share the data and resources of various workgroups? and

(3) How best to access various databases, archives or equipment resources which may reside inside or outside the organisation?

These questions must ultimately be addressed on both a site-specific and enterprise-wide basis. Through the 1980's, users began relying on local area networks (LANs) and "client-server" computing architectures to answer their requirements for resource sharing among users at individual sites. LAN usage in the general computing community has grown at a rate of 80% per year since 1985 [Pretty, 1991] and — with the commercial introduction of Unix-based GIS packages through the mid-1980's — LAN technology has been widely adopted in the GIS community [Miller, 1990].

Despite the popularity of LANs to connect users and resources within a single location, the GIS community has been justifiably reluctant to adopt *wide* area networks (WANs) to interconnect workgroups residing at different locations or to exchange data files with remote users, customers or suppliers. Due to relatively slow transmission rates and narrow bandwidths, bulk file transfers of GIS graphics and image data across dedicated telecommunication lines have been prohibitively expensive [Craig et al., 1991]. More importantly, long response times inherent in existing lower-speed networks have limited the practicality of carrying out routine enquiries and real-time display of data held in remote databases.

Where data delivery has been an issue, magnetic tapes and diskettes have traditionally been viewed as the media of choice for the distribution and/or exchange of data among GIS data producers and users [Crosswell, 1986]. A recent study indicates that Australian organisations involved in geoprocessing activities still transfer much of their spatial data files to remote users and/or outside customers in the same manner [Newton et al., 1992a].

On one hand, these off-line distribution arrangements provide a simple and relatively inexpensive means of delivering data from the supplier to an end user in a "reasonable" amount of time. Many programmes using GIS deal with facilities mapping or resource inventory data which may be 6-12 months old at best, so any minor delays in receiving the information can be easily accommodated.

Under such an approach, however, remote users of a corporate database cannot enjoy the same ad-hoc and on-going access to information enjoyed by their counterparts in head office unless they have replicated all relevant portions of the GIS database at their own site. Without such access, remote users may be less inclined to use the systems to compare phenomena observed in their particular area with those observed elsewhere.

In Australia, there are already examples of pressure being exerted by regional office staff to obtain complete replicas of their organisation's GIS database for use on their own system [Fox, 1992]. While optical disk storage certainly makes the efficient storage and transfer of even the largest GIS databases now technologically feasible, the potential duplications of effort, database inconsistencies and data backup issues all represent formidable management challenges once copies of the database leave the head office.

Finally, the update cycles on some program-driven GIS applications — particularly those involving cadastral mapping, navigational charts and real-time applications — can be considerably shorter than resource-based applications (e.g., days or weeks rather than months). While the database updates themselves are taking place, many of these programs must now keep out-of-date graphical databases and products in regional offices until more timely and efficient means of data communications and distribution can be implemented.

1.1.3 THE PROMISE OF HIGHER-SPEED NETWORKS

Introduction of higher-speed telecommunications systems may change the view of these organisations towards both the electronic transfer and overall management of such data. Important examples of new developments in data communications include:

- *Integrated Services Digital Network* – ISDN is now offered by telecommunications services around the world and provides customers with end-to-end data transmission rates of 64 K bits/sec and higher [Kirton, 1990];

- *Metropolitan Area Networks* (MANs) – Fibre-optic based packet switching services like SMDS in the United States and Telecom Australia's FASTPAC will provide high-speed (34 Mbit/sec) links between local area networks across and even between major cities, offering interface speeds comparable to those of the LANs themselves (10 Mbit/sec) [Pretty, 1991];.

- *FDDI* - Dedicated FDDI (Fiber Distributed Data Interface) networks already connect users within limited areas (e.g., across campuses or within metropolitan areas), transmitting data at rates up to 100 Mbit/sec.

- *Broadband ISDN* – B-ISDN promises to provide customers with 155 Mbit/sec access via optical fibre to a broadband network capable of supporting high-speed data and high-quality video services. Detailed standards for B-ISDN are expected in 1992 and commercial services are planned for Australia sometime after 1995 [Kirton, 1990].

Given these promised levels of performance, GIS users could potentially select and display large image and graphics files residing on archive systems across the city (or even the country) at speeds comparable to those where the data was stored on a local server. File transfers between sites could be handled between systems as quickly and easily as across a local area network. From a user's perspective, operating distributed data directory systems like NFS[1] across such networks may render it no longer immediately evident — or even necessary to know — where the files are physically located. In the longer term, the prospect of such seamless

[1] Network File System, from Sun Microsystems Ltd.

LAN interconnections may ultimately cause organisations to reassess the manner in which they acquire, distribute and manage their GIS-related data and system resources.

The rate at which organisations in the GIS community will ultimately adopt such technology remains uncertain, and many technical, operational and institutional factors must be considered in any corporate wide-area networking decision. However, unless any proposed solution offers: (1) fast file transfer capabilities; (2) high-speed access to remote hosts and (3) the seamless, transparent interconnection of local area networks in a client-server environment, it will probably be of limited value to GIS users in the organisation. With that in mind, GIS response time performance across a broadband network is an important indicator of whether or not organisations will adopt this technology and worthy of investigation.

1.2 RESEARCH HYPOTHESIS AND PROPOSED APPROACH

According to information provided by service vendors, new high-speed telecommunication services promise to provide users at remote sites with bandwidth and response-time characteristics similar to those enjoyed by local users. Based on this premise, it can be hypothesized that *broadband communications networks will provide the performance necessary to satisfactorily support the GIS application and data management requirements of a geographically-dispersed organisation from a single location.*

Ultimately accepting or rejecting this hypothesis will depend on the following:

(1) a clear understanding of a GIS and what constitutes GIS "application" and "data management" functions; and

(2) an objective assessment of what constitutes "satisfactory" performance in this regard.

The relevant terminology and concepts involved will be introduced and discussed in Chapter 2. Assuming that the average times required to complete a given operation on a stand-alone workstation (t_{ws}) or across a well-behaved local area network (t_{lan}) are deemed to be acceptable to most users, then the size of the incremental delay (Δt) encountered in carrying out the same operation across a broadband network should represent a reasonable measure of how efficiently that operation can be carried out across a metropolitan area network. To some extent, depending on the telecommunications tariff structure in place, it will also suggest how cost-effective it is to carry out such operations in that manner.

To test this hypothesis, defensible approaches will be developed to select representative logical network configurations and to identify representative data management and GIS-related operations which are frequently used or which place the highest loads on the networks and their components. Controlled testing will be undertaken to compare the time required to execute these operations on stand-alone workstations and across both local and metropolitan-area network links.

By examining how the performance times and data traffic patterns vary under different selected logical configurations and simulated traffic loads, the contribution of network delays to overall GIS response time may then be determined under different conditions. Using this information, the most desirable logical configurations (i.e., those which yield the least degradations in performance across the metropolitan area network) can then be identified.

1.3 IMPORTANCE AND CONTRIBUTION OF THE RESEARCH

1.3.1 Importance of the Research.

This research addresses an area which will become increasingly important to spatial data managers in future -- the anticipated performance of GIS over high-speed networks linking widely-scattered groups of users in different ways. As organisations complete the loading of their own GIS databases and as vendors continue to develop new tools for rapidly selecting and displaying imagery and graphics files, the demand for on-line access to these corporate "data libraries" from remote locations will become more intense.

Although local area network usage has grown at a rate of 80% per year since 1985 [Pretty, 1991], knowledge of comparative differences in GIS performance on stand-alone workstations versus across local and wide-area networks is still largely anecdotal. Implemented across metropolitan area networks, the Client-Server architecture will represent an important breakthrough for organisations with widely-distributed GIS workgroups wishing to make use of common equipment and data resources. However, until recently, relatively little documented testing was available concerning the dynamics and tradeoffs of GIS performance in a client/server environment.

The network monitoring proposed here provides comparative performance figures under controlled conditions. It identifies performance differences between alternative network and usage configurations, and examines the dynamics of data traffic across the net under different circumstances. Given the growing predominance of network-based computing, research in this area has the potential to identify GIS-related tasks or functions which could be optimised to operate more effectively in a client/server environment. This type of information is vital to: (a) organisations planning the extension of their communications services to include operational links to regional offices; (b) organisations interested in the distribution and/or dissemination of digitised mapping and remote sensing data to the general public; and (c) GIS software designers and vendors interested in satisfying customers with the above requirements.

1.3.2 CONTRIBUTION OF THIS RESEARCH

This particular research effort makes the following contributions to the overall field of knowledge in this area:

* The preliminary GIS usage monitoring work provides a defensible means of selecting the operations to be tested. Taken in a broader context, however, the approach provides the basis for a more rigorous and quantitative characterisation of GIS usage in an organisation than those simply relying on routine interviewing procedures.

* The research develops a practical but systematic approach to comparing GIS performance across metropolitan area networks under varying usage configurations and traffic loads.

* The results obtained from the performance monitoring provide statistically meaningful estimates of how performance changes as one performs the same GIS operations on a stand-alone workstation, across a local area network and between LANs across a metropolitan area network.

* The accompanying analysis of results adds to the body of knowledge concerning the dynamics of GIS processing and display performance across different client-server configurations, and provides a more complete indication of the effects of various GIS operations on network loading and performance.

* The research results and analyses are developed to give users a framework with which to evaluate the performance of different networking configuration options under specific conditions. Given the respective GIS usage profiles and the performance requirements at different sites, such a framework would assist in the optimal planning and implementation of wide-area network interconnections between widely-dispersed GIS workgroups.

1.4 LIMITATIONS OF THIS RESEARCH

The research described in this document was subject to several important limitations or extenuating circumstances. These are discussed in the following paragraphs.

Controlled environment: Except where indicated, the methodologies employed and results obtained in the performance comparisons were based on an environment of controlled system and network loading with no special tuning of operating system or network parameters. It is recognised that such conditions are rarely available in practice, and even minor changes in a Unix multi-user environment may have a major influence on performance in a shared database environment. Even so, the results should still provide a reasonable indication of relative performance differences at this stage of the FASTPAC service's development.

Hardware Actual performance results should be considered representative only of the particular combinations of hardware and network infrastructure components described in this document. The author accepts that hardware performance is increasing rapidly and that the results obtained on these tests may be different if completed at a later date on more up-to-date equipment. However, discussions concerning underlying behaviour and general network patterns identified in the research should have a somewhat longer half-life.

Software: Only the *Arc/Info* software (Rev. 5.01) was used in this testing. It is recognised that GIS software from other vendors — including Intergraph and Genasys II, for example — may exercise greater or lesser control over lower-level operating-system and data transmission functions. Further, many of the algorithms used in this version of *Arc/Info* may arguably take more i/o-intensive approach than comparable ones employed by other vendors [Healey, 1994].

Access to Broadband Services: Since Telecom Australia offers no commercial broadband data communication service in Tasmania, most of the research was completed using FASTPAC facilities in Melbourne and Clayton, Victoria. Because routine development and acceptance testing was already underway across these FASTPAC links, only limited time slots were made available to the research team. These slots included two ten-day periods: one in October, 1991 (across the FASTPAC test network in Melbourne) and the other in June, 1992 (across the operational Melbourne - Clayton FASTPAC link). No performance tests could be carried out across the FASTPAC link outside these periods.

1.5 ORGANISATION OF THE DISSERTATION

This dissertation is organised into seven chapters. In this first chapter, the research has been placed in context, the basic hypothesis and proposed approach have been described, and comments concerning the importance, contribution and limitations of the research were presented in turn.

The next three chapters introduce the major concepts forming the basis for this work and provide the reader with an appreciation of the rationale, planning and design considerations behind the experiments completed. In the second chapter, relevant background information concerning networks in spatial data management, data communications and performance evaluation is developed through discussion and literature review. Chapter 3 develops an approach to defining and quantifying GIS and network usage in an organisation, while Chapter 4 deals with experiment design and describes the equipment and procedures employed in the network monitoring and performance evaluation.

The last three chapters contain the results, analysis and discussion. The fifth chapter presents and compares the results of the various performance monitoring experiments and discusses their relevance with respect to user considerations of performance alone. Chapter 6 summarises the results and identifies possible topics for future research. Finally, Chapter 7 examines the impact of these results on GIS management and usage in an organisation and introduces some of the larger organisational and cost factors which must be considered in any information management and communications strategy.

2.0 BACKGROUND

Results from research described in subsequent chapters will be used to help determine whether broadband networks will really provide the performance required to enable organisations to manage their data and software resources from a single location. Before these results can be interpreted, however, some background information is required to appreciate: (a) the evolving conceptual views, market demands and operational requirements of spatial data networks; (b) the tasks and responsibilities associated with system and network management; and (c) the nature and demands of the performance-testing framework in which the experiments were carried out.

This chapter develops the relevant background by moving in from general concepts to specific project requirements. After defining some basic GIS and data communications terms in Section 2.1, the next section will describe how concepts of spatial information networks have developed from both the jurisdictional and organisational perspectives. In Section 2.3 , the author briefly introduces relevant concepts in to local and wide-area networking before moving into a description of distributed computing and the "client-server" architecture in Section 2.4.

Section 2.5 discusses the basic philosophies, options and considerations involved in GIS performance determination. After introducing general approaches to determining overall system and network performance, the author compares the objectives of this particular project with those of other GIS-related performance testing and modelling research completed or underway. The final section briefly summarises the goals and objectives of the GIS/Network performance-testing research which formed the basis for this dissertation.

2.1 BASIC TERMINOLOGY

In this dissertation, the term *geographic information system* (GIS) will be used in its broadest sense to include all types of computer-based systems used for the mapping, processing, analysis, management and display of spatially-related information. For the purposes of these discussions, the term GIS will be used to describe those systems devoted to mapping and civil engineering drawing, the query, management and display of geographically-related information, and the more complex spatial analysis and modelling [Antenucci et al., 1991].

Since the following data communication terms and acronyms are used extensively throughout this dissertation, brief definitions have been prepared and are presented in this section. Where required, further information describing the concepts introduced here will be presented in greater detail in subsequent sections.

While the functional distinction between the two is fast disappearing, *PC*'s are taken to mean any personal computer (IBM PC-compatible, Macintosh, NeXT, Amiga, etc.), while the term *workstation* is used to describe a higher-performance device with 32-bit processor, large amounts of main memory, and a high-resolution graphics display running the UNIX operating system.

A *local area network* (LAN) implies a network in which multiple workstations, terminals and peripheral devices are connected (physically or logically) to a single cable or shared medium [Pretty, 1992]. In practical terms, a LAN enables users within a limited geographical area (usually along a length of several hundred meters to a few kilometers of cable) to share output devices, storage units and more powerful processors. LANs typically move data at effective speeds equal to or greater than 2-3 megabits per second, with newer LAN technologies operating a speeds 10-50 times faster.

As the names imply, *metropolitan* and *wide area networks* (MANs and WANs) have wider geographical reach than LANs. Generally speaking, a MAN can cover a city while a WAN can cover a state, a country or even the world. The protocols and technology employed by these networks is different than that of a LAN in order to allow operation over greater distances. While the transmission speeds of such networks have been generally much slower than those possible across LANs, newer MAN and WAN technologies are offering speeds comparable to those across local area networks. Despite the differences in their reach, both types of networks have generally employed the same technology and have been applied to similar problems [Pretty, 1992]. Given the high costs involved in the installation and operation of the necessary equipment, most such networks are shared by multiple users.

Broadband networks are generally considered those which transmit data at speeds of 1 to 100 megabits per second or faster.

2.2 SPATIAL INFORMATION NETWORKS - PREDOMINANT VISIONS

The visions driving the development of spatial information networks today are not new, although many of the technological components required to implement such networks have only recently become available. This section examines how the vision of "networking" has developed in the GIS and land information management communities over the years and the contribution that broadband communications systems may make to these visions.

2.2.1 Integrated Mapping and Virtual Databases

By the late 1960's, the *integrated mapping* concept was being used to illustrate the registration of different "layers" or themes of spatially-related data to one another using a common spatial referencing framework (e.g., [Tomlinson, 1967]; [Roberts. 1968]; [McHarg, 1969]). By the early 1980's, this layers model was being widely used to illustrate the integration of data from different organisations within a jurisdiction. The term "network" is still sometimes employed euphemistically to indicate a heightened degree of inter-agency cooperation in the collection, management and sharing of spatially-related information -- in any form -- within or between two or more organisations.

Given the increased emphasis on end-user computing, these data sharing precepts evolved from the notions of centralised "land information databanks" common through the 1960's and early 1970's (e.g., [Cook et al., 1967]; [Roberts, 1968]) into the vision of building integrated land information networks. This vision conveyed the idea of linking together organisations responsible for the management of land-related information in a jurisdiction into a network to form a "virtual" geographic information system which could be queried in a manner similar to a single database. Hearle [1962] suggested such a concept at the state- and local government level almost thirty years ago, and both Palmer [1984] and Coleman [1988] provide more in-depth examinations of the institutional and technological issues involved. Finally, authors like Bell [1988] and Greenwell [1992] have provided more recent examples of how jurisdiction-wide land information networks have been implemented in the land information management community.

While the technology is now largely in place, the protracted and expensive database loading process and (especially) the lack of incentives for cooperation within some bureaucracies have limited the number of multi-participant, jurisdiction-wide land information networks operating today. However, given the more focused operating requirements, shorter decision cycles and quicker paybacks in a corporate environment a substantial amount of intra- and inter-office networking is taking place in large organisations. These *enterprise implementations* will be discussed in the next section.

2.2.2 Enterprise Implementations

Especially in Australia and Canada, the concepts and implementation examples cited earlier were based on jurisdiction-wide efforts rather than on the needs of individual organisations. However, the notion of "information as a corporate resource" [Diebold, 1979] and the subsequent development of *information resources management* practices promoted through the early 1980's encouraged many individual organisations to begin reassessing the manner in which they managed their various hardcopy and computer-based data holdings. "Enterprise Implementation" efforts often encompassed one or more of the following actions:

(1) identifying both the data holdings of common interest to several workgroups across an organisation *and* the unit(s) within the organisation responsible for the collection, verification, management and update of each type of data;

(2) modelling the flow and use of various corporate data by different workgroups and individuals within the organisation;

(3) defining an agreed-upon level of standardisation of selected data types in order to reduce the duplication of effort involved in their collection and management;

(4) systematic planning, modelling and implementation of the individual database management systems which would manage this corporate information;

(5) developing the procedures, database management facilities and communications infrastructure required to integrate these various data holdings and (where applicable) interconnect the databases involved.

The basic objectives and activities involved in such efforts were similar in nature to the more jurisdiction-wide concepts discussed earlier. In fact, some important components of such systems (e.g., coordinate system definition and base mapping) may have been adapted from wider government efforts.

Because of their size and more restricted mandates, individual organisations can reach consensus on project definition, justify the expenses involved, implement their respective data collection and/or integration efforts and achieve results much more quickly than those involved in multi-participant efforts covering an entire jurisdiction. Not surprisingly, large utilities and municipalities involved in enterprise implementations were among the first to identify the operational requirements to integrate smaller GIS- and facilities management-related databases on stand-alone systems with larger corporate databases residing on mainframes [Popko, 1988].

2.2.3 Spatial Data Infrastructure

The *information infrastructure* concept has contributed to spatial information networking discussions at the professional, corporate and jurisdiction-wide levels for over ten years [Branscomb, 1982]. Proponents of this concept envision a spatial data infrastructure analogous to networks of federal highways or electric power grids in countries around the world. As with the power grid analogy, the physical location of the information source itself would be less important to the end user than its continuing availability, reliability and cost [Anderson, 1990].

The concept of "plugging in" to a wide range of standardised information sources was first introduced during the 1960's (e.g., [Goldmark, 1972]). Like "networking", the term "infrastructure" was at first employed in a limited sense to refer only to integrated— but independent — data holdings. For example, the notion of treating certain classes of spatial information as "infrastructure" *per se* was developed in [Hamilton et al., 1974]. While the term "information infrastructure" had since been coined to refer collectively to the various media, carriers and even physical infrastructure used for information delivery [Branscomb, 1982], the term "intelligent infrastructure" was also becoming commonly used by the mid-1980's to describe structured data holdings in the automated mapping/facilities management community [Robinson, 1986].

By the late 1980's, however, the term was being recognised in a much broader context, and the notion of infrastructure as an *enabling agent* (i.e., enabling users to "plug in" to independent databases) was adopted once again. Anderson [1990] suggested this type of infrastructure should possess the following three important characteristics:

(1) the contents (data), conduit (telecommunications network) and flow-control procedures should be standardised;

(2) the major sources and users must be networked together; and

(3) the network must be customised for easy third-party access.

Especially since the late-1980's, the concept has been proposed in support of accelerating geographic information exchange standards efforts, selected national mapping programs and the establishment of nation-wide spatial information networks in the United States [Mapping Sciences Committee, 1993], the United Kingdom [Rhind, 1992] and Canada [McLaughlin, 1991].

2.2.4 Characterising Current GIS Networking Objectives and Capabilities

Especially since 1986, "networking" in the land information management community has moved from euphemism to reality. Rather than simply indicating the presence of formal or informal agreements between organisations for streamlined data sharing, communication networks are actually being established in order to increase the speed and volume of information between or (especially) within organisations. Networking issues within organisations have been discussed at length by authors from major utilities and municipalities (e.g., [Ingoldsby, 1991]). Recent examples of provincial or state-wide "jurisdictional" networking efforts can be found in both Australia ([Greenwell, 1992], [SLIC, 1992]) and Canada [Forrest, 1992].

With so many networking efforts proceeding now underway at both the enterprise and jurisdiction-wide levels, one group's concept of a "network" can be very different from another's. To understand and compare different networking efforts, Keen [1991] suggested a framework for describing existing or proposed data communication capabilities within and between organisations. Table 2.1 (from [Zwart et al., 1992]) summarises the generic data communications functions of such networking efforts, and indicates the relevant capabilities commonly found in the spatial information community in 1991.

Many communications services now provide electronic mail and file transfer capabilities to widely-dispersed users, although customers are still constrained by incompatibilities between various public and private networks. Providing access to remote applications and resources is now considered a basic requirement within local area networks and is becoming generally available across metropolitan-area networks as well. Access to corporate information systems by internal staff is now commonplace, and some databases are now available to outside users in a jurisdiction on a cost-shared or unit-charge basis.

Operating examples dealing with genuine cross-linked access to multiple databases are much harder to find, since such programmes require a degree of agreement on common data standards and operating procedures not commonly present between different organisations. While many projects are now in the planning stages (e.g., [Loukes et al., 1990]), very few are actually in operation. Multi-participant, jurisdiction-wide examples such as those now being implemented on the Island of Guam [Starling et al., 1991] and in the province of Alberta, Canada [Langille, 1991] are notable exceptions.

Table 2.1

Electronic Data Communications Capabilities Generally Found in Different Organisational Settings
(from [Zwart and Coleman, 1991])

Function / Area of "Reach"	Standard Messages (E-Mail)	File Transfer Between Systems	Remote Login; Single- Database Transactions/Inquiries	Multiple, Cross-Linked Database Transactions
Within a Single Workgroup	✓	✓	✓	O
Within a Single Location	✓	✓	✓	O
Same Organisation; Different Locations	✓	✓	✓	O
Multi-Participant Efforts Users with Same Hardware & Software	✓	✓	✓	O
Multi-Participant Efforts: Users with Different Hardware & Software	✓	O	O	-
Anyone, Anywhere	-	-	-	-

LEGEND

✓	O	-
Capabilities available; Widely used	Few examples outside major corporations or research networks	Uncommon; Not commercially available

Factors Facilitating Broadband Network Implementation

Most land information networks currently in operation in Australia have been concerned with transferring textual data from a dedicated host computer to remote terminals. LOTS in South Australia is one example of this [Sedunary, 1988]. Only recently have some jurisdiction-wide networks – such as the Western Australia Land Information Access Network [Bennett et al., 1988] and New

South Wales' Land Information Hub [SLIC, 1992] begun transferring large volumes of vector and image data as well.

Given this interest in networking by land information management organisations and the large data volumes involved, the spatial data handling community would seem to potentially represent a significant market for broadband communications services. Senior Telecom Australia staff believed so and, in 1990, they commissioned an investigation into the characteristics and communication requirements of selected users in the GIS, land information systems (LIS), automated mapping/facilities management (AM/FM) and computer-assisted drafting (CAD) sectors [Newton et al., 1990].

In that study, the authors identified the following factors contributing to the increased interest in broadband telecommunication technology in the spatial data handling community:

(1) *Growth in Data Volumes:* Raster scanning of subdivision plans and digitizing of cadastral data will dramatically increase the size of LIS databases previously limited to textual data. (A typical property map image file can be 16 Megabytes in size.) In remote sensing, receipt of up to 1200 Gigabytes of satellite imagery data *per day* is anticipated as users deal with increased resolution of satellite imagery, a growing number of satellites in use and more organisations offering the data itself.

(2) *Growth in Data Sharing and Distribution*: Utilising data collected from different sources in order to reduce duplication of effort and generate new information products has been the *raison d'etre* of corporate land information management efforts for the past fifteen years. The current development of *land information directories* in various jurisdictions around the world is evidence of market demands for concise information describing contents of spatial data held by various organisations. AUSLIG's on-line Land Search

Directory, the National Resource Information Centre's (NRIC) FINDAR system, and the Environmental Resources Information Network (ERIN) all represent operational or proposed examples of such directories in Australia.

(3) *Increasing Demand for Real-time Systems:* Hardware improvements have brought about dramatic improvements in processing speeds, memory, storage, and graphics capabilities over the past 30 years. In effect, these new capabilities have altered *user expectations* regarding minimum "acceptable" response times for particular computer-based operations. To help identify candidate organisations most likely to benefit from adoption of broadband communications technology, the investigators classified applications in the LIS/GIS/Remote Sensing and CAD communities according to two criteria:

- *Perishability of the data which the organisation maintains* to support the application or programme; and
- *Level of need for interactive processing or display* over MANs, WANs or global networks

Table 2.2 classifies the organisations and summarises the applications under review according to these criteria.

(4) *Internationalisation of Industry:* The development of global communications networks (cable plus satellite) has been a facilitator for the internationalisation of industry. Growth in the business of international couriers underlines the needs of organisations to transfer large volumes of documents between countries. Nissan Australia currently utilises such services to transfer tapes containing CAD drawings and specifications between Japan and Australia, but is exploring electronic transfer options combining high-speed ground-based and satellite communication services [Newton et al., 1990].

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="2">Data Perishability</td></tr>
<tr><td>HIGH</td><td>LOW</td></tr>
<tr><td rowspan="2">Need for Interactive Processing or Display over Metropolitan-Wide-Area or Global Networks</td><td>HIGH</td><td>Type I Organisations

Existing applications requiring high-speed data communications capabilities

e.g.
Weather Monitoring & Forecasting
National Defence
Remote Surveillance
Emergency Despatch
Vehicle Fleet Management</td><td>Type III Organisations

Most GIS applications currently fall into this category.

e.g.
Regional Planning
Conservation & Environment
Municipal Planning & Engineering
Forest Inventory & Management
Soil Surveys & Monitoring</td></tr>
<tr><td>LOW</td><td>Type II Organisations

Real-time processing carried out in-house on LANs. Data transmission to remote sites on tape or diskette via courier.

These organisations likely to migrate to Type I status, but first require sufficient justification for distributed analyses and fast file transfer to remote centres.

e.g.
Medical Imaging</td><td>Type IV Organisations

Little or no change over time in the information content of data supporting these particular disciplines or applications.

Coverage of relevant areas can be stored on optical disk for use in the office or in the field.

e.g.
Exploration Geology</td></tr>
</table>

Table 2.2:

Two-Factor Classification of Differing Organisational Demand for High Speed Communication Networks
[Newton et al., 1990]

(5) *Integration of Telematic Activities and Applications:* As telecommunications move into an era which permit integration of voice, text, data and images, applications are emerging to take advantage of such network capabilities. Examples include linking text documentation and specifications with CAD drawings, linkage of subdivision lot plans with owner details on title documents, and combination of image, text and possibly voice communication for GIS consultative planning applications.

Because their operational requirements can be clearly defined and since their database implementation timetables are usually much more contained, broadband communications efforts may well be justified earlier by individual organisations than by those managing state-wide, multi-participant networks. Demands from users to be able to display and/or download large graphics or image files stored on remote systems are already being met in the current operations of, for example, the Sydney Water Board's IFIS network [Chapman, 1991] and the Victoria Department of Conservation and Environment's GIS Section (Alexander et al., 1992]. More important, it is being taken into account in the current planning efforts of local government organisations like the Melbourne Information Technology Services, state organisations like the New South Wales Roads and Traffic Authority [MacDonald, 1992] and the Commonwealth Government's Environmental Resources Information Network [ERIN, 1991].

To summarise, the predominant visions and euphemisms associated with networking in the spatial data handling community have dealt to date primarily with the integration and sharing of information within corporate and/or and jurisdictional boundaries. To date, most users in this community have been preoccupied with the more fundamental issues of system interconnection, database and system security, and data exchange between different systems. However, the levels of both interest and user expectation with respect to GIS performance in a networked environment are finally beginning to increase for the following reasons:

(1) corporate attention is finally turning from initial database loading to data management and distribution within and between organizations;

(2) the standards for LAN performance, compatibility and interoperability developed and demonstrated over the past ten years have turned client/server computing into a paradigm for office automation within the business community; and

(3) the increasing integration of large graphics, image and now video files in the GIS environment can only increase data traffic across networks.

Before examining the management responsibilities and actual performance testing requirements, the next section offers a closer look at the physical realities, capabilities and future applications of local, metropolitan and wide-area networks.

2.3 INTRODUCTION TO LOCAL AND WIDE-AREA NETWORKING

Until the mid-1970's, the predominant data communication systems in most business environments were relatively simple in concept and dealt with moving data between central systems and remote terminals. By 1985, however, the increasing emphasis on distributed data processing techniques (i.e, connecting *computers* in different geographic locations to one another) rendered the data communication systems much more complex and problematic [Martin, 1988].

The overall benefits of computer networks have been discussed at length by Champine et al., 1980], Martin [1988] and Gunton [1989], among others. By allowing computers to communicate with one another, networking permits users to share access to scarce equipment resources (e.g., printers, plotters, databases, etc.), makes possible communication applications like electronic mail and file transfer between sites, and enables distributed processing [Pretty, 1992]. As well, today's networking technology enables users to expand their facilities with some degree of vendor independence and to incorporate special purpose processors, storage units or input/output devices as required.

While Section 2.2 discussed the driving philosophies which have provided the impetus for high-level interest in networking within the spatial data handling community, this section introduces the practical and technical details. In order to understand the performance and behaviour of GIS operations in a networked environment, it is useful to first briefly review some basic concepts in data communications technology. After introducing the important notions of network architecture and communications protocols, this section will briefly discuss the characteristics of local and metropolitan and wide area networking from an end-user's perspective.

2.3.1 Network Architecture and Protocol Suites

2.3.1.1 BASIC DEFINITIONS

A *computer network architecture* is a set of functions, interfaces and protocols which enables devices to communicate with one another on-line [Palmer, 1984]. The architecture is composed of a number of modular functions layered such that — while each layer is designed to operate independently — higher-level operations are built on functions provided by the lower layers [Chorafas, 1980].

Generally speaking, *communication layers* are designed to create error-free links between the physical channels which connect the computers, while *networking layers* use these links to create "virtual circuits" which direct data from the transmitter to the receiver. *Application layers* use this communication path to control I/O devices, access files and transmit data from application programs

[Chorafas, 1980]. Two examples of such layered models will be discussed later in the next section.

Network Protocols are formal sets of rules or specifications for coding messages exchanged between two communication processes on a network [Voelcker, 1986]. Protocols govern data control and format across a network, and a variety of protocols exist to ensure these communications are conducted effectively.

2.3.1.2 PROTOCOL SUITES

Software for networking is often viewed as a stack of layers (consisting, in turn, of specific protocols) based on the model developed by the International Standards Organization (ISO). Two different "stacks" of layered approaches are in common operation today [Baker, 1992]:

- *OSI* — A 7-layer model was developed by the International Standards Organisation (ISO) and has become known as the *Open System Interconnect* (OSI) suite of protocols. (See Figure 2.1)
- *TCP/IP* — A 4-layer model comprising the Transmission Control Protocol/Internet Protocol (or TCP/IP) suite of protocols was originally developed for the ARPAnet research network in the United States and funded by the U.S. Defence Department and various research organisations.

Originally developed by computer user groups and European telephone companies, the OSI model helped unify world telephony and provided a clear framework and explanation of the functions required for computer communications. However, it is regarded by some as being too cumbersome for high-speed networks since, in a typical transmission — each data packet must pass down through all the layers of the sending computer, up and down through

Layers 1, 2 and 3 of all forwarding computers and finally up again through all seven layers of the destination computer [Wittie, 1991]. (See Figure 2.2.)

USER APPLICATION PROGRAM	
LAYER #7 Provides communications services in formats comprehensible to the application programs.	***APPLICATIONS***
LAYER #6 Transforms data to and from various standard formats.	***PRESENTATION***
LAYER #5 Synchronizes and manages the exchange of data among senders and recipients.	***SESSION***
LAYER #4 Provides transparent and error-free transfer of data between non-adjacent network nodes.	***TRANSPORT***
LAYER #3 Performs message routing for data transfer between non-adjacent network nodes.	***NETWORK***
LAYER #2 Manages network access and performs error detection for messages moved between adjacent nodes.	***DATA LINK***
LAYER #1 Specifies the electrical connections to the network and transfers messages between adjacent nodes.	***PHYSICAL***

Figure 2.1

The Open Systems Interconnect (OSI) Model

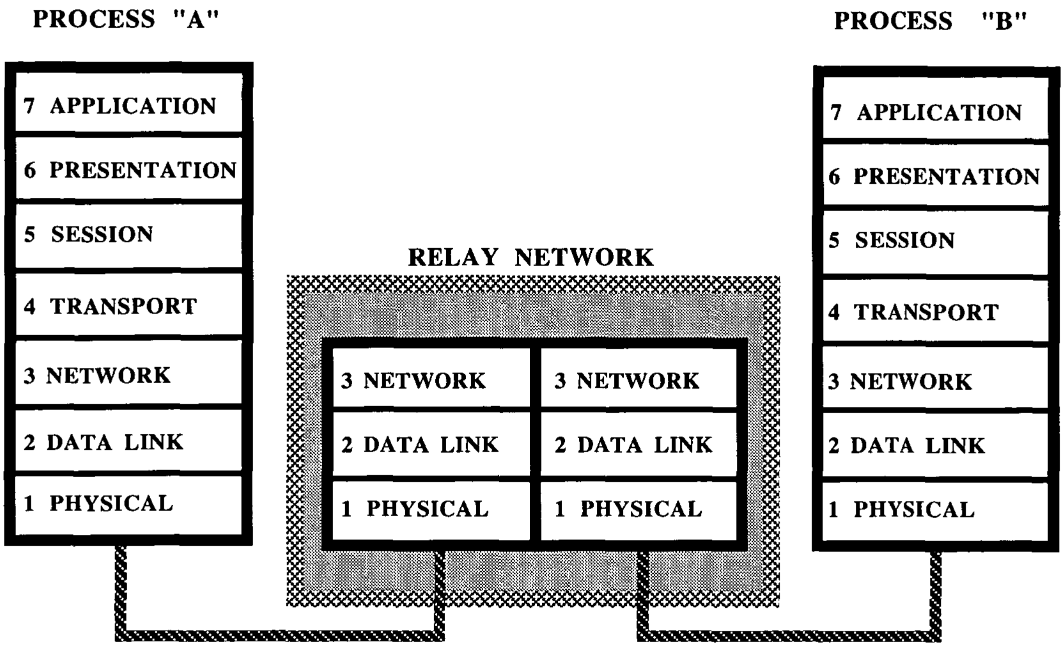


Figure 2.2

Communication Through an OSI Network

By comparison, the TCP/IP suite of protocols became a *de facto* standard by the early 1980's as a result of its early use in the implementation of the U.S. Defense Department-funded Internet. Its longer-term popularity was secured through subsequent bundling with the 1983 release of Berkeley Unix 4.2 (itself a de facto standard operating system for university workstation networks). Consequently, although the OSI model has been officially sanctioned by most international organisations, the TCP/IP suite of protocols remains in much more common use and, if anything, is moving from a *de facto* to official networking standard in some jurisdictions [Baker, 1992].

2.3.1.3 THE TCP/IP PROTOCOLS

The TCP/IP collection of networking protocols was developed primarily for and in the UNIX community. As shown in Figure 2.3, the TCP/IP model consists of four layers. Above the network hardware interface is the IP layer which handles internetworking (i.e., the splitting of data packets for network transmission and their reassembly at their destination).

The TCP and UDP protocols sit atop the IP layer, and application programs use one or the other of these to ship data across the network. TCP was designed to provide reliable, sequenced delivery of packets over (relatively) long-lived network connections. In fact, it requires a log-in connection between the sending and destination workstations. UDP, on the other hand, is more of a "no frills" protocol which sends packets to a remote host, but makes no assurances regarding their delivery or the order in which they arrive [Stern, 1991]. UDP is better suited for "connectionless" communication environments like NFS (i.e., where no explicit log-in is made) in which no context is required to send packets to a remote machine. This difference will prove to be significant in subsequent experiments.

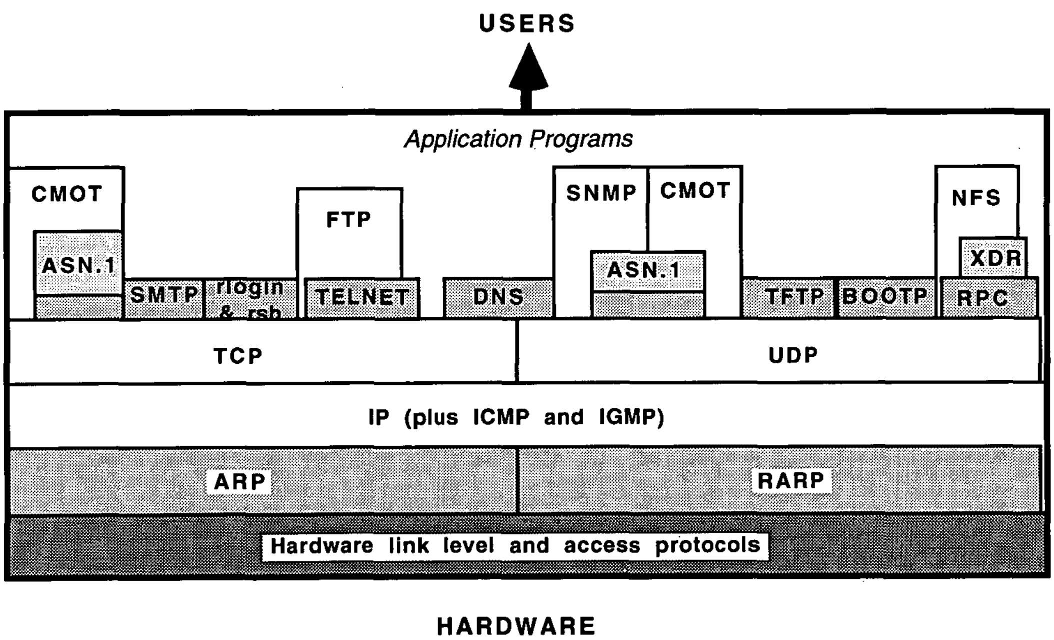


Figure 2.3
Higher-level TCP/IP Protocol Dependencies
(Baker, 1992)

A number of application programs with their own unique protocols sit atop TCP/IP. These programs provide UNIX users with standard tools for dealing with file transfer (FTP), electronic mail and terminal emulation (TELNET). Another application sitting atop TCP/IP — the Network File System (NFS) — provides on-line shared file access to remote directories in a transparent and integrated manner. NFS will be discussed further in Section 2.4.3.

Strictly speaking, the TCP/IP protocols do not precisely fit into the more general OSI model. However, the functions performed by each OSI layer do correspond to the functions of each part of the TCP/IP protocol suite and provide a good framework for visualising the respective relationships between the various protocols [Stern, 1991]. (See Table 2.3.)

Table 2.3

Relationship between TCP/IP and the ISO Model

OSI Layer Number	OSI Layer Name	Corresponding TCP/IP Function / Service
7	Application	NFS and NIS
6	Presentation	XDR
5	Session	RPC
4	Transport	TCP or UDP
3	Network	IP
2 1	Data Link Physical	Ethernet

2.3.2 Local Area Networks and LAN Bridging

2.3.2.1 LOCAL AREA NETWORKS

A *local area network* (or LAN) connects together (physically or logically) multiple workstations, terminals and peripheral devices to a single cable or shared medium [Pretty, 1992]. (See Figure 2.4.)

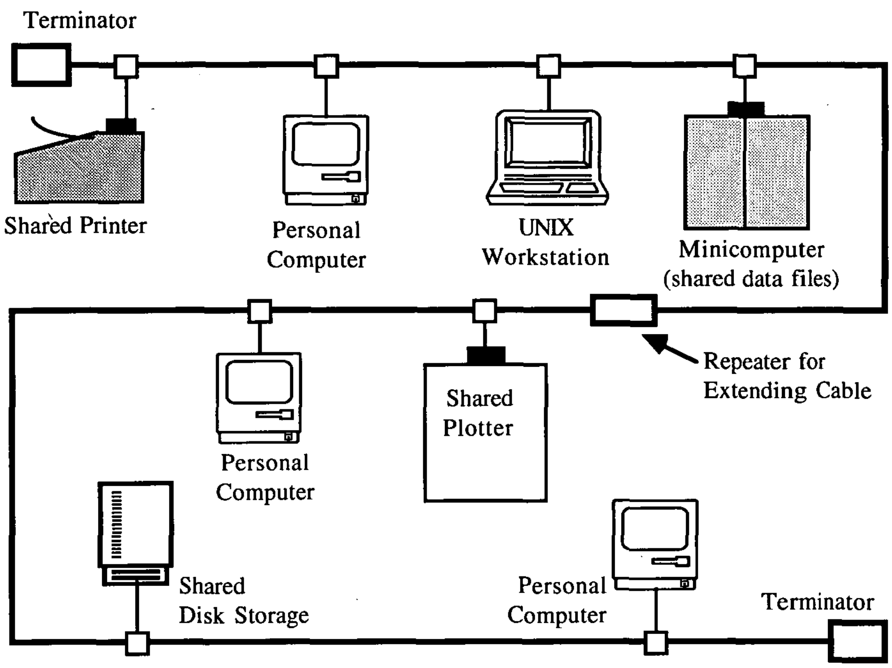


Figure 2.4
Simplified Ethernet Local Area Network

During the 1980's, over 100,000 LANs were set up in offices and laboratories around the world to link workstations to printers, share files and send electronic mail [Wittie, 1991]. LAN usage in the general computing community has grown at a rate of 80% per year since 1985 [Pretty, 1991], with networks now extending into schools, libraries, laboratories and offices around the world using telephone lines, optical fibres and satellite links.

Several accepted and standardised types of LAN technology now share the market, including:

- IEEE-802.3 (CSMA/CD or *Ethernet*)
- IEEE 802-4 (*Token Bus*)
- IEEE 802-5 (*Token Ring*)
- ANSI FDDI (Fibre Distributed Data Interface)

All IEEE standard LANs operate using a physical or logical broadcast bus (often called a *shared medium*), which nowadays may be copper cable, optical fibre or even airspace using wireless signal. Medium Access Control (MAC) layer packets or frames called *Datagrams* are broadcast onto the shared medium where all network stations can observe them. These datagrams contain both Source and Destination Address identifiers which uniquely identify, respectively, the station which transmits the packet and the station which should receive the packet.

LAN stations are "promiscuous listeners", i.e., they observe all packet traffic travelling across the LAN. The station will read the Destination Address within each datagram and — if it does not match its own address, it is ignored. However, when a station recognises its own Destination Address within a datagram, it reads in that frame and passes its contents up to the next protocol layer for processing [Telecom Australia, 1992].

To date, the Ethernet and Token Ring technologies have tended to dominate the market [Wittie, 1991]. While the FDDI technology is relatively new and is much less tolerant of line breaks or faulty stations, it does provide data at much higher speeds than its older counterparts [Pretty, 1992]. Since Ethernet LAN technology is being used in these experiments, this will be discussed in detail.

Ethernet Local Area Networks

The first Ethernet LAN was built at Xerox Corporations Palo Alto Research Centre in 1975 to connect various computers within the Centre by coaxial cable. Originally designed at MIT [Metcalfe et al., 1976], it proved so successful that Xerox (in conjunction with Intel Corp. and Digital Equipment Corp.) published specifications for a 10 Mbps Ethernet in 1980 and started licensing companies to produce the required interfaces.

Ethernet generally uses coaxial cable (fibre and wireless versions are available as well) to connect all participating computers onto a single broadcast bus. Packets transmitted onto this bus are seen by all stations, and each station must examine the destination address information contained in the packet header to see if it is "one of theirs".

Ethernet LANs typically move data at effective speeds equal to or greater than 2-3 megabits per second, with newer LAN technologies operating a speeds 10-50 times faster [Clarkson, 1993]. Controlling the sending and receiving of data by different users on the Ethernet is governed by a medium access control algorithm called *CSMA/CD* (or "Carrier Sense, Multi-Access with Collision Detection"). Wittie [1991] suggests that CSMA/CD makes Ethernet work "like an old-fashioned telephone party-line with courteous users." All stations listen to the cable (carrier sense) and, if it is idle, any machine can transmit (multi-access). If more than one station happens to transmit during the same period, their messages destroy each other in a collision on the cable. When a collision occurs, all stations detect the resulting noise, temporarily cease their transmissions and wait a random amount of time before trying again [Pretty, 1992]. While few collisions occur during periods of low network usage, they may occur frequently in cases where the network is heavily loaded with small packets.

2.3.2.2 LAN INTERCONNECTION

Since LAN performance can diminish as one increases the number of stations connected, a single LAN may have to be broken into multiple LANs to maintain performance as traffic increases. More to the point here, since LANs have a limited geographic reach, considerable effort has been expended in finding effective ways of connecting LANs across wide areas.

Multiple LANs may be interconnected to form a larger network using a variety of packet relaying devices, including MAC bridges, routers and gateways [Pretty, 1992]. *Bridges* examine the destination addresses of each datagram to determine if that packet should be sent on to the next LAN or if it is destined for a local station. If the packet is being sent to the next LAN, the bridge forwards it along; otherwise it ignores it. Because of this, a bridge must be connected to at least two LANs and know (or be able to learn) the respective station addresses on each of the LANs involved.

A number of LANs connected by bridges is called an *extended LAN*, and these extended LANs can usually support a greater number of users and reach over a longer distance than a single LAN. In this environment, bridges provide three important functions [*ibid*]:

(1) *Traffic Segmentation* —they keep overall traffic levels low by isolating work group traffic within individual LANs;

(2) *Remote Bridging* — they allow LANs to be geographically separated from one another so long as long-distance transmission paths are available between bridges; and

(3) *LAN Internetworking* — they can interconnect LANs of different types, although this type of work is usually handled by gateways.

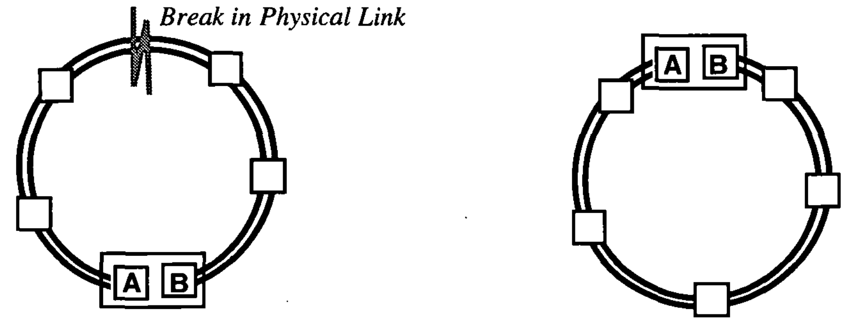
Routers and *gateways* (the terms are often used interchangeably) provide a higher level of functionality than bridges [*ibid*]. Like bridges, they examine additional information contained within the packets (particularly structured addresses) to forward the packet to the appropriate destination. However, since these addresses can also contain ownership information, they can ensure increased transmission security as well.

2.3.3 Metropolitan Area Networking and the FASTPAC Service

Metropolitan and Wide Area Networks are designed to cover a wider geographical area than LANs. While MANs were originally intended to cover metropolitan areas while WANs were designed to reach across a country or around the world [Pretty, 1992], the technologies now employed in such systems are now generally the same and the terms are fast becoming interchangeable. The protocols and technology employed by these networks is different than that of a LAN in order to allow operation over greater distances. While the transmission speeds of such networks have been generally much slower than those possible across LANs, newer MAN and WAN technologies now offer speeds comparable to those across local area networks.

At least officially, the Metropolitan Area Network (MAN) technology designed so that users can share network resources over a metropolitan area approximately 50 Km. in diameter has been standardised by IEEE Committee 802.6. This MAN standard, known as Distributed Queue Dual Bus (DQDB), was designed to provide high-speed packet-switched LAN interconnection and fast file transfers at access rates up to about 45 Megabits per second. Unlike other high-speed wide-area networking services in use in the early 1990's, the IEEE 802.6 standard offers both higher performance and longer geographical reach using shared network resources and bandwidth.

The *Distributed Queue* part of the name refers to the unique means by which the nodes on the loop gain access to the shared medium. The DQDB technology employs two separate unidirectional buses running in opposite directions past nodes (hence the "*Dual Bus*" portion of the name). To provide increased security and fault tolerance, subnetworks are typically configured as loops. (See Figure 2.5.) In the event of a physical break somewhere in the link, the network is logically reconfigured by closing the buses at the co-located head ends (Boxes A and B in Figure 2.5), thereby re-forming the dual bus architecture and minimising any interruptions in service.



Loop in Initial Configuration **Loop Reconfigured Around Break**

Figure 2.5

Fault Tolerance in the IEEE 802.6 DQDB MAN/WAN

FASTPAC is a metropolitan area networking service now being introduced by Telecom Australia which is based on the DQDB standard. The service employs an IEEE-specified "Queue Arbitrated" medium access method which allows bandwidth to be dynamically and statistically shared among all connected stations, thereby providing a LAN-like datagram service across the interconnected networks [Telecom Australia, 1992]. With these capabilities, FASTPAC provides a transparent LAN bridging service at the Medium Access Control Layer. In effect, stations on interconnected local area networks act as though they are all part of the same LAN.

Telecom is offering two classes of FASTPAC access, supplying peak data transfers at both 2 and 10 Megabits per second (Mbps). The FASTPAC 10 service relies on the existing and planned fibre optic transmission network to provide its reach within and between cities. The FASTPAC 2 service employs a copper cabling transmission system which links the customer's site to the nearest FASTPAC node.

In the major capital cities in Australia, 140 Mbps DQDB loops connect FASTPAC exchanges within the central business districts. At present, the existing FASTPAC 10 MANs in Melbourne, Sydney, Canberra and Brisbane are being connected to those in each of the other centres using 34 Mbps long-distance trunk lines. FASTPAC 2 services are planned for Adelaide and Perth by 1993 (with others to follow), and services will be upgraded to FASTPAC 10 if and when demand warrants [Montgomery, 1992].

While originally intended as a *metropolitan* area networking service, much of the initial demand for FASTPAC 10 services is coming from customers interested in interconnecting LANs located in different cities — traditionally a wide area networking application. While appreciating the differences in formal IEEE definitions, Telecom Australia FASTPAC staff have effectively begun to use the terms *wide-area networking* to encompass FASTPAC, ISDN and lower-speed services. This will be the case in this thesis as well.

2.3.3.1 FUTURE METROPOLITAN AND WIDE AREA NETWORKS

Optical fibre transmission systems are really capable of accommodating much higher speeds and data volumes. With this in mind, the International Consultative Committee on Telegraph and Telephone (CCITT) is presently developing standards for a Broadband ISDN (B-ISDN) service which would handle voice, data, image and video traffic simultaneously. The central element of B-ISDN is the "asynchronous transfer mode" (or ATM) switch, a cell relay device able to supply channels operating at speeds of 155 and 622 Mbit/sec accesses [Kirton, 1990]. Already in pilot testing, ATM technology will provide higher performance and greater bandwidth efficiency than the current generation of broadband networks, greater reliability and availability of service, and more adaptable quality of service to providers of public and private networking services [Smith, 1992].

Clearly, the thousand-fold increase in capacity over today's "high-speed" services offered by B-ISDN should provide the throughput necessary to accommodate high-bandwidth services like high-definition television, video-on-demand, and real-time video-conferencing [Tonkin et al., 1992]. However, there is no clear agreement as to either: (a) whether the market demand -- if it exists -- will be driven by new mass-market applications (like HDTV) or existing improvements to large corporate internetworking services; (b) how soon value-added services will be able to take advantage of this capacity; or even (c) whether or not the current tariffing philosophies can accommodate the service-integration flexibility offered by B-ISDN [Finnie, 1992].

Section 2.3 has introduced some of the relevant terminology, important concepts and performance behaviour capabilities inherent in local, metropolitan and wide-area networking. These technical details — like the driving philosophies discussed in Section 2.2 — collectively represent a second "piece of the puzzle" in obtaining an overall appreciation of the considerations involved in testing GIS applications in a networked environment.

While a high-level understanding of the technical components is desirable, it is very important to understand how networks are being employed by GIS users during the early 1990's. While the networking technology provides the medium for communications, it is the models of distributed computing — and the *client/server architecture* in particular — that have actually redefined the notion of end-user computing in the workplace. The next section introduces this topic and describes some of the potential implications of the client/server computing with respect to GIS performance.

2.4 DISTRIBUTED COMPUTING AND THE CLIENT-SERVER ARCHITECTURE

2.4.1 The Concept of Distributed Computing

The traditional centralised computing environment can be characterised by small numbers of large-scale mini- or mainframe computers, with shared storage devices attached via hardware I/O (input/output) channels and multiple users connected via terminals possessing varying levels of on-board "intelligence" or processing power [Katz, 1991]. Such systems have formed the basis for most major data processing and information systems applications since the 1950's, and most major commercial GIS packages were originally developed for use in either mainframe or minicomputer environments.

Especially since 1988, there has been a major shift away from such centralised environments towards a more distributed model of computing and data processing. The term *distributed computing environment* implies a situation where — rather than consolidated into a single shared processor — the processing tasks and data are distributed among separate components connected by a network. Terminals, graphics workstations, disk systems, CPU's and various input and output devices are all attached to the network to build the overall computer system, with all the various components capable of being accessed in a relatively transparent manner. The major ingredients of distributed systems are [Dall et al., 1988]:

(1) *Families* of computing units;

(2) *Localisation* of processing to appropriate positions or functions;

(3) *Interconnection* of processing units for the access, movement, sharing and/or distribution of data;

(4) *Movement Management* of the data, transactions, requests and operational aspects.

With careful planning and management, distributed computing returns responsibility for data integrity and access to local users and refocuses individual transaction loads back to their natural locations.

At the operational level, distributed systems generally offer faster response time, greater availability, a higher degree of system security, less complexity and, in many cases, lower-cost computing solutions than more traditional mainframe or minicomputer solutions. At a higher level, the notion of distributed computing often provides a better fit to the complex structures and often multidisciplinary nature of modern organisations and offers greater user involvement in information management activities. This has strong appeal to end-users and unit managers who want more control over (and faster, more secure access to) their data *and* to some data processing managers eager to reduce the workload on their larger systems.

One of the most popular approaches to distributed computing employs the *client/server model*, which is discussed in the following section.

2.4.2 The "Client-Server" Architecture

2.4.2.1 BASIC COMPONENTS AND CONCEPTS

The shift from centralised to distributed computing in today's marketplace has been due primarily to the rapid development and acceptance of the *Client/Server* model of computing, and most distributed computing today is based on this type of architecture [Katz, 1991]. As the name implies, a group of workstations (clients) relies on one or more servers residing elsewhere on the network for access to data files, application software and, in certain cases, more powerful computing services. In such an environment, the data retrieval aspects of a database query can largely be carried out independently of the data processing and display tasks.

In the client/server model, client workstations make *open, close, read* or *write* file requests of other designated machines on the network. The machines, usually called *file servers*, are processors with large disks capable of storing large amounts of data for other users on the network. Such servers are generally optimised to provide high-speed retrieval of large data or database files from disk. [Miller, 1990].

Stand-alone workstations are those which contain all the software, data and on-board resources required to perform any desired GIS operation. By comparison, *dataless workstations* are those which may possess a minimal amount of disk storage to run the operating system and maintain basic working space, but rely on file servers for storage of the requisite data and application software. Finally, *diskless workstations* are those which may possess large amounts of internal memory and are capable of fast, independent processing, but which rely on external servers for both temporary work space and storage of all operating system, application software and data files.

2.4.2.2 DATA FLOW IN THE CLIENT-SERVER MODEL

In a typical operation in which the user wishes to read a file stored on a remote server, the Client workstation will issue a data *read* request. After travelling from main memory to the Client's network interface board, the request will travel across the network to the address of the specified Server. After entering through that unit's network interface board, the component messages within the request are directed to the Server's main memory.

If the requested data already resides in the Server's cache memory, they are sent directly to the Client along the reverse path. If (more likely) the data are not in memory, the file system will issue a request to retrieve the data from the disk (or whatever I/O device is specified). The requisite data are retrieved into memory along the internal, high-speed backplane bus and the overall response is sent out across the network to the Client workstation (via the respective interface boards).

Clearly, the overall response time of such a process depends on the individual performance of the storage, processing and communications components involved in the system. While improvements in all three of these technologies have all made key contributions, it has been the introduction of remote file management services like the Sun Microsystems' Network File System (NFS) and others which finally provided the transparency required for distributed computing. Exactly how this development at the operating system level improves end-user network computing is discussed in the following section.

2.4.3 The Network File System (NFS)

Remote file management services are a collection of tools and processes embedded within the operating system which enable users to transparently access, view and manipulate files and directories stored on remote disks in the same manner as they would if they were stored on their own machine [Stern, 1991]. Remote file system implementations like Sun's Network File System (or "NFS') on UNIX workstations, PC-NFS, *Novell Netware* and *Banyan Vines* (on PC-compatibles) and *Appleshare* (under System 7 on Macintoshes) have become extremely popular and common in many organisations. Given this popularity and since — as will be seen later — the behaviour of NFS has a measurable impact on overall performance, this section provides a brief introduction to the objectives and characteristics of network file systems.

The UNIX implementation of NFS is still the most ubiquitous file system model. In this implementation, a file is simply an uninterrupted stream of bytes named within a hierarchical naming system based on files within directories within still larger directories, and so on. The data structure of the file system contains all the information required to locate and keep track of the various components of the file, which may be physically distributed in various locations across the disk.

The major innovation of Sun's NFS is its ability to map remote file systems into the directory structure of the local machine [Katz, 1991]. From the user's perspective, NFS is almost invisible: the file names themselves give no indication of whether they are stored locally or on a remote machine [Baker, 1992]. The theoretical basis for this "virtual file system" is described in [Sandberg, 1985].

Whenever an application makes a *read* or *write* request, UNIX calls the file system to handle the request. Access to the remote machine is handled via a synchronous "remote procedure call" (or RPC) which initiates a procedure on the remote machine. Because this RPC is synchronous, the Client machine must wait until the server has completed the call and returned either the requested data or status acknowledgement [Stern, 1991].

The *NFS Protocol* is a collection of procedure calls, descriptions and parameters built on top of a Remote Procedure Call. By design, NFS is a *stateless* protocol: to ensure simple recovery from system/network crashes, each procedure call is self-describing and the server keeps no track of past requests. If the server (or a network connection) does go down, the client workstation simply keeps re-sending its request until it receives an answer or times out.

Since no complex recovery processing is required, such a system is ideal from data management and crash recovery perspectives. However, the performance demands of stateless systems are much higher than those of "stateful" systems where both servers and clients keep track of all requests in real-time. In the stateless case, the server (i.e, the unit writing data to a remote machine) must ensure that all modified data and metadata is committed from cache memory onto remote disk before returning control to the user. After sending a single packet of data, the server must wait for an acknowledgement of receipt from the remote machine before sending the next packet of data in the overall transmission. The cumulative effect of these incremental delays on overall performance can become quite significant, especially in transactions involving larger files and many hundreds or thousands of incremental writes to the remote disk [Katz, 1991].

The effects of NFS-induced delays on GIS performance in a client-server environment will be discussed in Chapter 5. For further information, Katz [1991] provides an excellent overview of NFS in the larger context of overall network performance. Sandberg [1985] provides a conceptual overview of the early design and implementation issues, while Stern [1991] provides an in-depth description and analysis of NFS characteristics, performance specifications and tuning considerations.

2.4.4 Discussion

2.4.4.1 PERFORMANCE RELATIONSHIPS IN A CLIENT/SERVER ENVIRONMENT

As described in Section 2.4.2, there can be a long instruction path associated with processing a network-based I/O request. Instructions, data and acknowledgement packets may flow back and forth between various disk, memory, backplane and network components many times during a typical transaction involving a large file. Depending on the hardware and software present in a given installation, the overall architecture may or may not be optimised to permit fast processing between network and disk interfaces [Katz, 1991]. Moreover, as implied above, the speed of NFS itself is coupled to the performance of the disk system employed [Rosenblum et al., 1992].

These inter-relationships have been recognised, and technology developments in data processing, storage and communications are all improving performance in a client/server environment. However, these improvements are being introduced into systems at different rates. For example, processor performance (rated in MIPS) has roughly doubled every 18-24 months over the past ten years. While not evolving as quickly, newer networking technologies like ISDN, FDDI and DQDB (described in Section 2.3.3) now offer an order-of magnitude improvement of data throughput along the backbone from the early, 3 Mbit/sec Ethernet products offered in the early 1980's [Wittie, 1991]. However, due to the capital costs, logistics and time involved in installing new cabling throughout a building and/or across a country, many organisations are still employing networks with capacities of only 10 -16 Mbit/sec.

While storage capacities of magnetic disks have been doubling once every three years, mechanical limitations are preventing further improvements in the speeds of actually seeking and returning data (the random I/O rates) on single disk units. *Disk array* units are now addressing this data access problem by replacing a few large-format disks with a very large number of smaller-format ones, thereby putting more units to work on retrieving the same large data request [Katz et al., 1989].

While individual improvements are being made on all three fronts, the *collective* effects of improvements to storage, communications and processing performance in many network applications are still not well understood [Pretty, 1992]. While *file transfer* performance across networks has been measured or modelled by various researchers (e.g., Bachmann et al., 1989], [Fenwick, 1990] and [Yang et al., 1992]), results of end-to-end performance simulation and measurement research *within client-server environments* have only recently begun appearing in the literature [Cabrera et al., 1991].

2.4.4.2 MANAGING RESOURCES IN A CLIENT/SERVER ENVIRONMENT

In a pure distributed storage model, since datafull workstations may be clients for some applications and servers for others, it is the respective locations of *the user, the data* and *the host processor* which determine the logical network configuration in use at any particular time. However, since commercial GIS or DBMS software and databases can consume large amounts of disk storage, many operations have placed these resources on a limited number of (often) centralised file servers. By employing high-performance workstations linked to high-capacity file servers, both vendors and users alike hope that the client/server model can offer the advantages of both distributed processing and centralised storage [Katz, 1991].

Some believe this trend represents a return to more traditional configurations distributed processing was meant to replace. Clearly, the client/server model does possess the most well-known weakness of more centralised approaches in that the failure of a single server or the network backbone often affects all connected client workstations on the network. Further, depending on the types of client workstations employed and usage conditions involved, network capacity may represent a significant bottleneck to overall system performance even where high-speed servers are in use.

A growing number of users believe these concerns are balanced off by operational advantages. For example, since individual users in a client/server environment see the same file system regardless of the machine they are using, the view of data storage and access to network resources is pervasive and transparent. As well, from a management perspective, it is usually much easier to administer a centralised system, manage software updates and provide archival storage services under such configurations.

Conflicting Concerns of Different User-Groups: the "Project vs. Programme" Framework

In some organisations, technical arguments often have little to do with any final strategic decisions concerning centralisation or distribution of system and data resources. However, the nature of the predominant workload in an organisation can significantly influence attitudes towards the management of system and data resources. In this context, "predominant workload" can be classified in terms of whether it is *project-* or *programme-driven* and whether it covers the *short-* or *long-term*. (See Table 2.4 for examples.)

Table 2.4

Classification of Workload in an Organisation

	Project-Driven	**Programme-Driven**
Short-Term	Educational Assignments Short-term scientific research projects	
Long-Term	Large Engineering Projects (e.g., Boston Harbour Cleanup)	Forest Inventory GIS Property Valuation Database

Both Edwards et al. [1990] and Gunton [1989] suggest that many opponents or detractors of centralised system management may be present in organisations where the workload is focused on short-term projects. In such cases, the end-users are often researchers or managers operating on stand-alone PC's who have taught themselves the necessary data processing and GIS skills required to complete their own particular task. Since there may be little apparent requirement for data continuity, these people see little immediate need for the overhead of data documentation, well-maintained and labelled data libraries, and well-defined lifecycle management of hardware and software systems.

In such environments, if the project data is ever required at a future date, the end-users usually retrieve the data from the project files themselves. Since the frequency of such retrievals cannot be predicted and because such users often place little or no value on their own time (even though they are often well-paid and the times involved may be substantial), they still may see little economic justification in anything more than ad-hoc labelling and archiving procedures geared to their own requirements.

Supporters of centralised network and data management tend to operate in program-driven environments where more formalised data processing policies and routines are considered an essential means to an end. In organisations which rely on regularly-changing information contained in large databases, such considerations as data integrity, consistent database update and secure access to system and data resources are often vital to on-going operations. In such situations, the contribution of the GIS to the overall programme goals and service levels is well understood and the value of a well-maintained database — or, at least, the cost of taking actions based on incorrect or out-dated information — can often be quantified and measured. Not surprisingly, centralised management of data resources in such cases is viewed as an essential means of recouping (or at least *sustaining*) large investments in data and equipment ([Chandler, 1989]; [Evans, 1987]).

Regardless of the nature of the workload in an organisation, formalised management of network and data resources need no longer carry the negative connotations present in earlier data processing environments. As Edwards et al. [1990] suggests, modern networks make the concept of centralised data management more philosophical than physical. Under current client-server architectures, the data archives, management software and management personnel may all reside at different locations throughout the network (i.e., rather than at a "central" location). In short, the information management environment can be organised in response to the operational needs of the organisation rather than vice-versa.

The performance of a particular application or group of applications across a network is a crucial consideration for network planners and managers alike. In some cases, such issues may influence the type of configurations and equipment eventually adopted. Issues and concerns involved in system and network performance determination will be introduced in the next section.

2.5 PERFORMANCE DETERMINATION IN A NETWORKED GIS ENVIRONMENT

The Need for Hard Numbers

A recent study carried out for the Government of Canada [IDON, 1990] suggested that the technologies required to satisfy the networking demands of federal government GIS users may already exist in the marketplace. However, designing and building *cost-effective* GIS networks within and between federal government organisations would require defensible answers to six important questions, including:

- What kind of information is to be moved?
- How large are the data volumes involved?
- How quickly do the data files need to travel from sender to receiver?
- How often and/or how regularly will data be sent?
- How far the data was being transmitted; and
- How much are people prepared to spend to build and keep such a network in place?

Even in cases where headway is being made in developing common standards and institutional agreements, the inability of many jurisdictions to predict future usage and to reliably quantify the benefit/cost tradeoffs involved has hampered the physical establishment of state- and federal spatial information networks [IDON, 1990]. Rigorous, defensible approaches to predicting the potential nature and volume of GIS-related communications network traffic will be required to support such an effort. This is as true for single organisations as it is for multi-participant projects.

This requirement for "hard numbers" reinforces the need to systematically and defensibly determine and/or predict the performance of a particular application or group of applications across a network. Clearly, numerous authors have suggested the fundamental importance of performance determination within the overall framework of the system life cycle and structured system and database design processes (e.g., [Ferrari, 1983], [Heidelberger et al., 1984], [Stonebreaker, 1985], and [Jain, 1991] among others). However, while performance analysis has formed an important component of many GIS selection processes since at least the early 1980's (e.g., [Tomlinson et al, 1981]), most of the procedures and results of such evaluations have remained unavailable due to commercial confidentiality constraints and competitive pressures.

Researchers have been developing more systematic and rigorous approaches to the determination of GIS performance on stand-alone configurations since the mid-1980's (e.g., [Marble et al., 1986]). However, with the exception of recent investigations at the University of Edinburgh (e.g., [Sloan et al., 1992]), relatively little has been done to rigorously examine how GIS performance may change in a client-server environment across local or wide area networks. It is this requirement that represents the central focus of this thesis.

This section examines the role and benefits of performance determination within the larger context of the system life cycle process and discusses the problems and considerations involved in obtaining reliable, meaningful information on performance. After describing the general philosophies, objectives and components of the process, the contributions of other relevant research efforts and investigations into measuring GIS performance in both stand-alone and client-server environments are examined.

2.5.1 Components of the Overall Performance Evaluation Process

The evaluation of system performance is a fundamental component of the system life cycle process [Heidelberger et al., 1984]. It provides greater insight into hardware and software behaviour under different conditions, and can provide the information necessary to identify and determine the location, scope and effects of system bottlenecks. When applied carefully and objectively, the performance evaluation process can provide a quantitative means of:

(1) comparing different systems and algorithms;

(2) measuring the effects of various conditions or phenomena on overall system performance; and

(3) determining the relative impacts of incremental system adjustments, extensions and "improvements" [Wagner, 1992].

As well, it provides the user with the information necessary to utilise hardware and software more efficiently *and* predict future resource requirements more reliably [Dowers et al., 1990].

A comprehensive evaluation process must look at the basic functionality of the system, the fundamental correctness of the results and the efficiency of performance. Most formal evaluation processes are broken down into three components, including:

(1) *basic functionality*

(2) *results verification* and

(3) *performance measurement.*

Examining basic functionality of a system confirms whether or not a system possesses the range of functions necessary to address a particular application and deliver the result in the form required by the user. The verification process establishes that — given valid and legitimate input — the system will consistently deliver correct and complete results [Berg et al., 1982]. The measurement process determines how quickly and efficiently this operation or series of operations can be performed.

2.5.1.1 PERFORMANCE VERIFICATION

To be completely effective, verification testing must examine the entire "problem space" of a system [Wagner, 1991]. Given the scope of functionality, input data and operating options available in most DBMS and GIS packages today, the verification process can be problematic at best. As a result, systems are typically tested using an extensive subset of the combinations of operations and data sets which could be employed by most users. Such tests, while extensive and costly in themselves, usually fall far short of simulating all possible combinations.

The issues of basic GIS functionality and verification *per se* are outside the scope of this thesis. Rather, it is the measurement and comparison of GIS performance across stand-alone and (especially) networked configurations that forms part of the basis for this research. Before moving directly into GIS-related performance measurement issues and examples, the following section discusses the overall objective of performance measurement and the practical alternatives available to the investigator when attempting to measure system performance.

2.5.2 Performance Measurement

Performance measurement seeks to establish the relationship between the size of the workload generated by a particular problem and the costs of the system resources (i.e., hardware, software and labour) consumed in the process. Wagner [1992] defines the relationship between workload and system resource requirements as the *Performance Space*, and it is the goal of performance measurement to investigate and define the bounds of this space

In the experiments conducted for this research (to be described in Chapters 4 and 5), individual response times for selected GIS operations were measured under controlled conditions in a variety of stand-alone and networked configurations. While (it will be argued) this is a valid approach to assessing system performance, there are other means of determining performance determination as well. This section will introduce the three best known approaches to assessing system performance and present the reasons why the direct-measurement approach was eventually selected. In addition to describing the methods themselves, the section will also discuss the factors and parameters which must be taken into account in order to produce reliable and repeatable results.

2.5.2.1 MODELLING, SIMULATION AND DIRECT MEASUREMENT

The performance space of a system may be analysed and assessed by either predicting, calculating or observing the system responses for specific processes or to controlled phenomena. The three basic approaches to performance measurement include:

(1) Analytical Modelling

(2) Simulation Modelling

(3) Direct Measurement

Analytical and simulation modelling both require a model of the system which will embody and reliably represent its behaviour under varying circumstances. Such a model should contain parameters representing factors which may be varied to portray different systems. Values of these parameters can depict the amount of services demanded by the customers and the rate at which they are handled by the system. Ideally, the model can use this information to determine contention among system resources or processes and the resulting effect it may have on performance throughout the system. As a result, one can gain a deeper understanding of the how the system performs and can study its behaviour in a controlled fashion [Jain, 1991].

Both approaches employ models to determine the systems behaviour. *Analytical modelling* represents the system by a series of mathematical equations. By assigning specific values to the various parameters, the equations can be solved to obtain performance measures which estimate how the system behaves under the conditions specified. *Simulation* uses a computer program which acts like the system. This program keeps track of the contention for resources represented in the model, and then calculates the system's performance based on what it has observed [McNair et al., 1985].

While they often demand less time and resources than direct measurement, analytical modelling and simulation efforts usually require details concerning the specific algorithm(s) employed, hardware characteristics and overall system implementation. Even then, the results of such predictive efforts may be inaccurate and unreliable under certain conditions [Jain, 1991]. Provided a prototype or operational system already exists, the simplest and most direct approach to performance measurement is by directly observing system responses to specific operations under controlled conditions.

2.5.2.2 ALTERNATIVE APPROACHES TO MEASURING PERFORMANCE

Four different approaches are commonly proposed as alternatives for measuring system performance. In order of complexity, these include:

(1) Determining the bounds of performance space by locating points on either side of the boundaries defining that space;

(2) Determining the relationship(s) defining the performance space through statistical analysis of test observations conducted under different conditions;

(3) Defining the limits of system performance under one or more "average-case" scenarios and assuming that performance under all future operating conditions will usually fall at or near these limits;

(4) Defining the limits of system performance under "worst-case" scenarios and assuming that performance under all other conditions will therefore fall within acceptable limits;

Systematic Approaches

Wagner [1991] refers to the first and second approaches listed above as "systematic approaches" to the problem of performance measurement. In both cases, after identifying potential influences on performance (e.g., data complexity, file size, number of individual points, orientation or arrangement of data, etc.), a rigorous approach is taken to determining the sensitivity of system performance to individual or combined variations in these various parameters.

Both these approaches (the first one in particular) possess the ability to create a general overall picture of the performance space. Even in situations where the characterisation may be incomplete, investigators can statistically predict system performance for a given set of conditions and forecast the resource demands of future applications. On the other hand, substantial time and effort may be required in taking all the observations necessary, and even then these may yield reliable characterisations of only a small number of operations. A considerable degree of skill is required in selecting or creating the most meaningful indicators of performance to be used in such tests [Wagner, 1991].

Application-Specific Testing

The third and fourth alternatives — sometimes referred to as the "application-specific" approaches — make use of problematic, real-world operations or applications, with either "average" or "worst-case" characteristics used as test cases to determine system performance. In cases where the system in question (or a similar one) is already in operation at a comparable location, "average" or representative characteristics may be determined through monitoring of system usage over a period of time. By comparison, the "worst-case" scenarios may qualify as such due to especially-unusual data characteristics, processing demands or stringent and pre-defined output requirements of a given application.

Application-specific approaches possess some important limitations which must be noted [Hawthorn, 1985]. First and foremost, the underlying assumption that the system can handle anything less severe than the pre-defined worst case may simply not be true in all cases. Even if it is, the results may provide little information concerning the nature or extent of the performance space itself. Since performance characteristics of individual component operations may vary significantly, little can be said about performance under varying conditions. Finally, "average" and "worst-case" scenarios can change as system usage evolves: the results of a "worst-case" scenario are valid only as long as no *new* worst cases are subsequently identified for that particular application.

Even with these limitations, however, the systematic approaches can still be appealing in both theoretical and practical applications. For example, maxima of particular variables are often employed in statistical modelling and simulation as a means of both predicting system performance and identifying weak links or bottlenecks in a given process [Heidelberger et al., 1984]. Further, they possess the advantage of measuring system performance under at least some semblance of contemporary operating conditions [Stonebreaker, 1985].

Finally, while they only deliver a limited amount of information, the suite of observations required take far less time to complete than those necessary for the systematic approaches described earlier. This is particularly important in situations where time, equipment and/or manpower resources are limited or where there are only short windows of testing time available.

Given this rationale, the basic demands of the research and the time constraints involved, an application-specific approach was selected for the performance testing. (See Chapter 4 for details of the experiment design.) The following section introduces some of the key generic parameters involved in direct measurement under such an approach and briefly describes relevant selections influencing the follow-on experiments.

2.5.2.3 PARAMETERS AFFECTING DIRECT MEASUREMENT

Direct observation and measurement of computer systems performance usually requires specification and control of two or more of the following parameters (extended from [Jain, 1991] and [Wagner, 1991]), including:

(1) Performance Metrics

(2) Performance Monitors

(3) Workloads

(4) Factors and Levels

(5) Command Scripts

Performance Metrics are the specific system or resource indicators being observed. They are usually expressed as either a *direct measurement* or as an *indirect statistic.*

Provided the qualifiers are well-documented and understood, *response time* is one particular measure of performance easily recognised by end-users. Mayhew [1992] defined system response time as being "...the time between user input and *the onset* of system output". By adding a separate component *system display time* (defined as "...the time required to complete the display of the system response on the screen once the system has begun to do so"), she acknowledged the measurable influence of separate on-board processors to speed up graphics display.

In this particular research effort, response time (in fact, *differences* in response time) was selected the principal performance metric. For the purposes of these experiments, *response time* shall be defined as the sum of the system response and system display times for a given operation. This will be calculated using UNIX-based monitors which record the clock time at the start of each successive operation in a command script. These figures will then be objectively verified using a stopwatch.

Performance Monitors are devices used to record changes in the system during the testing period. Monitors may be completely external to the system (e.g., a stopwatch to measure elapsed response time) or may be implemented internally (using hardware, operating system utilities or special-purpose software).

Provided the monitors employed and the system or network under test are synchronised, the performance indicators and resource requirements of the system can be determined at various times throughout the test. However, the process of observing something can actually change it; this is especially true with software-driven monitors, whose operation may typically consume measurable amounts of system resources on their own.

Workloads are the jobs run on the system for testing purposes, and may be real or synthetic. *Real workloads* allow the system to be measured under "real-world" conditions, they are usually unique, are rarely extendible and do not offer the investigator any degree of control or variability over key performance parameters.

By comparison, *synthetic workloads* are artificially designed and generated to provide the investigator with much greater control over data and functions which may influence system performance. By varying the combinations involved, investigators may examine different points inside and outside the performance space. However, since these synthetic workloads rarely resemble those found in practice, relating experimental results to normal conditions may be difficult.

Factors are variables of the workload which may or may not affect system performance (e.g., density, orientation and relative complexity of the features in the dataset). *Levels* are the respective values these factors may be assigned for different tests. *Factor/Level Combinations* refer to the various possible combinations of factors and levels used in the actual testing. In simple designs, the optimum level for each factor is determined, any possible interaction between factors is ignored, and the chosen level is used throughout the evaluation process. At the other end of the spectrum, a "full factorial design" takes all possible combinations of factors and levels into account. Tests involving some subset of all possibilities are called "fractional factorial designs".

Given the potential combinations of variable parameters which may influence the results of a performance-testing experiment, it would be costly and time-consuming to attempt to define the performance space of all operations under all conditions (data, hardware, operating systems, etc.). The complexity increases considerably when the examination moves from a stand-alone workstation to a client/server configuration on a local area network, and increases even further when examining performance across a metropolitan area network. As the complexity increases, it becomes more and more time-consuming to rigorously examine all possible factors in a systematic manner.

Table 2.5 contains selected examples of key factors which may influence GIS performance across a metropolitan area network. Clearly, the most important dimensions of a particular application must be identified in order to select the most appropriate choice of "fixed" versus "variable" parameters for any subsequent sensitivity analysis.

Finally, *Command Scripts* are batch-invoked programs, macro commands or scripts which actually issue the commands and operate the monitors. Under normal conditions, many system operations would be performed interactively. However, in repeated experiments which try to exclusively assess *system* performance, inconsistent human responses may cause significant variations in elapsed time. Where one or more operations are being repeated several times, running them as a batch process minimises any variation due to inconsistencies in operator interaction.

Table 2.5

Technical Factors Affecting GIS Response Across a Network

GROUP	FACTOR
MAN-Backbone	Bandwidth Capacity
	Avg. vs. Peak Bandwidth Loading
	Traffic Management Protocols Employed
	Number and nature of Gateways, bridges and routers on the MAN
LAN-Related Hardware and Software	Speed of Network Controller
	I/O Throughput Capacity and Data Transfer Rate of the LAN Server
	Average Percentage Utilisation of the LAN Server Capacity
	Protocol Implementation
	Packet Size
	Mix of Disk*full* vs. Disk*less* Workstations in the LAN configuration
LAN-Backbone	Nature of Interconnect devices (e.g., bridges, routers, servers)
	Bandwidth Capacity
	Avg. vs. Peak Bandwidth Loading
	LAN Topology
	Length of Cable
	No. of hosts per cable
User Workstation	Nature of W/S — i.e., Diskless vs. Diskfull W/S
	I/O Throughput Characteristics (i.e., capacity & data transfer rate) of the workstation disk
	Amount of Available Workstation Memory
	Memory-Resource Utilisation (i.e., use of RAM disk)
	Workstation CPU Performance
Application Software	Proprietary Data Structure
	Algorithms Employed
	Processing Optimisation Techniques Employed
	Memory-Management Techniques Employed (e.g., Swap-Space Requirements)
	Propensity of application software to fragment data files
Data	No. of Polygons
	No. of Arcs
	No. of Coordinate pairs in each arc
	Overall Size & Composition of Data Files
	Size and structure of attribute database
Operations & Logistics	On-going Mix of Real-Time vs. Bulk File Transfer Applications across the LAN
	Nature of commonly-employed user activities
	Concurrent vs. sequential running of batch procedures
	Capability of using *multiple* disks rather than a single disk for processing.

2.5.3 Relevant Research on GIS Performance Testing

2.5.3.1 EARLY EFFORTS

Given the phenomenal growth of GIS software acquisition, applications and research since 1980, it is surprising that relatively little beyond anecdotal descriptions has been published concerning the procedures and results of GIS performance evaluations. Calkins [1983] pointed out this lack of available information and stressed that performance testing should be considered an important component of the system development process. Goodchild et al. [1986] echoed this by claiming that, while software exists which apparently can fulfil many customers' functional GIS requirements, the lack of information concerning the performance of such software makes it difficult to predict future resource needs.

In examining previous efforts, Stefanovic et al. [1987] suggested that a more systematic approach to benchmarking of computer mapping systems was required and claimed that the operating environments employed in application-specific GIS benchmarks could not adequately replicate real-world conditions. Peuquet et al. [1990] also expressed concerns, indicating that this continuing lack of reliable GIS performance information can hamper the assessment of potential contributions offered by incorporating new technologies (i.e., solid-state disks and parallel processors) into GIS production operations.

A number of theoretical studies *have* been undertaken over the past fifteen years which examined the efficiency of various GIS operations. While the analysis and comparison of alternative approaches to polygon overlay has been especially well-documented (e.g., [Lam, 1977], [Aronson, 1982], [Guevara, 1983], [Christie, 1984], and [Wagner, 1988], among others), the computational complexity of such operations as point triangulation and spatial searches have also been examined in considerable depth [Howes, 1991] and [Yang, 1992]). While all these certainly add to the overall knowledge concerning the performance of specific operations, very few efforts have been documented which take a more synoptic view of GIS performance.

In 1980, Roger Tomlinson and Associates Ltd. completed and documented an application specific benchmark for a forest inventory GIS completed for the Saskatchewan Department of Natural Resources [Tomlinson et al., 1981]. In that particular study, a number of different computer mapping/ GIS software packages were tested using a collection of functional requirements previously defined as being necessary to the customer's application.

This particular report represents a milestone as the first published paper to present a documented example of comprehensive testing and comparison of GIS software performance against a list of stated user requirements. Even so, while this work is widely recognised as being a good first attempt, there was no attempt to verify the correctness of the system output. Moreover, since the application-specific tests employed were far from complete or systematic, the bounds of the performance space were not determined [Wagner, 1991].

Many organisations have since employed some level of systematic performance testing (or "benchmarking") during the GIS selection process. However, for reasons of client confidentiality, proprietary information or previous agreements between the customers and competing vendors, very little information was ever publicly released concerning the procedures involved and results of these benchmarks. While application-specific methodologies *were* developed and heuristics concerning reliable GIS performance metrics *were* obtained, they generally remained within the realm of GIS consultants and system vendors. The trends towards systematic performance measurement and testing — already well-entrenched in the larger data processing community (e.g., [Ferrari, 1978] and others) — had not yet been adopted by the GIS consulting or research communities.

2.5.3.2 RECENT PERFORMANCE TESTING RESEARCH ACTIVITIES

Since the mid-1980's, at least six different research groups have recognised the need for the development and documentation of rigorous and repeatable approaches to GIS performance evaluation. Four of these efforts, discussing performance measurement and sensitivity analysis in a stand-alone and minicomputer environment, are discussed in this section.

Goodchild et al. — University of Western Ontario, Canada

Goodchild et al. [1986] published the results of an early attempt at developing a general model of the system acquisition and benchmarking process for use under actual production conditions. By defining the component sub-tasks involved in selected GIS operations, varying the pertinent characteristics of representative data sets (e.g., number of polygons/sheet), and then measuring system performance on a limited number of factor/level combinations, this model was used to predict system performance and resource requirements. While the empirical test results were not really suitable for defining the entire performance space or for comparing different software packages, the authors' research did offer a systematic approach to predicting the demands that known volumes of work would place on a given system.

Marble, Wagner et al. — The Ohio State University

During roughly the same period as Goodchild's efforts, Marble et al.[1986] proposed the adoption of systematic approaches developed originally for use in relational database management testing [DeWitt, 1985]. In a follow-on effort, Marble et al. [1989] compared the performance of selected operations on one particular GIS package (*Arc/Info 4.01*) on two different hardware platforms. In both efforts, Marble and his colleagues examined the performance issue largely from an end-user's perspective: the GIS was treated as a "black-box" in which the user had no access to source code and could not modify or enhance the hardware, operating system or application software in any way.

While most of the tests in [Marble et al., 1989] were application-specific, the performance of at least one operation was compared systematically on different datasets of varying size and density. Although the results were reasonably well-documented, the lack of information concerning the comparative characteristics of each platform makes it difficult to determine the relative contributions of (e.g.) CPU speed power, disk speed and available memory to overall performance. As well, the measure of performance used to present the results (a ratio value) cannot be extended, rendering it problematic to extend and/or compare this work with any similar research [Hawke, 1991a].

More recently, the work of Wagner [1991] must be recognised for extending the extending the existing research and — most significantly — for developing a more formal and comprehensive conceptual framework for the systematic evaluation of geographic information systems performance.

Amundsen -- University of Hawaii

Unlike Marble et al.'s "black box" approach, Amundsen [1989] worked with access to the full source code of a raster-based GIS (*OSU MAP-for-the-PC*) in order to examine the effects of hardware on performance. Multiple tests were run using gridded data sets of varying resolution, and software-based performance monitors were incorporated into the GIS source code to examine the effects of different processors on response time. While the results may not have been statistically reliable (due to relatively low repetition rates), the results did yield a general indication of the performance-time differences in selected operations on PC/XT vs. PC/AT microcomputers.

Hawke — University of Auckland

Working in a vector environment (*ArcInfo 5.01*) and assuming a "black-box" view of the system, Hawke [1991a] provides a well-documented and systematic examination of the sensitivity of GIS response times of two specific operations (topology construction and polygon overlay) to varying host processors and dataset sizes. This particular research was subsequently extended to examine the performance sensitivity to data *complexity* as well [Hawke, 1991b]. Chan [1991] from Edinburgh also investigated the effects of data point distribution and complexity on GIS performance, but chose a very different metric for use in the comparisons.

It can be argued that the particular functions chosen by Hawke and others for testing (i.e., usually involving topology creation or polygon overlay) represent "worst-case", processing-intensive operations which would not be frequently invoked by the majority of users in more "mature" GIS environments. While this may be true (as discussed in the next chapter), the documentation and systematic approach adopted by Hawke represents a healthy trend towards more rigorous and controlled testing procedures, better planning and selection of relevant performance metrics, and more intelligent presentation of testing results.

2.5.3.3 PERFORMANCE TESTING IN A LOCAL AREA NETWORK ENVIRONMENT

Although there is finally evidence of growing interest in GIS performance testing on stand-alone systems, very little has been published concerning the role and performance of such systems in a client-server environment across local- and wide area networks. While some vendors have published articles in GIS periodicals and newsletters (e.g., [Miller, 1990], [Camarata, 1992]), most of these have simply introduced the concepts and described the various networking options now open to institutional users.

Early examples of GIS performance testing in a client-server environment are now appearing in print. Dowers et. al [1990], Healey et al. [1991] and Gittings et al. [1991] from the University of Edinburgh provide some of the earliest documented examples of application-specific GIS performance testing in a networked environment. As part of the same research effort, Lau [1991] examined the influence of hardware, operating system and networking characteristics on GIS performance. While much of this group's efforts have been specifically directed towards hardware influences, their results have contributed to the general knowledge in three important areas:

(1) The research acknowledges and makes a first attempt at isolating the various classes of factors influencing GIS performance in a client-server environment (i.e., *hardware*, *network*, *operating system*, *application software* and *data*) [Sloan et al., 1992];

(2) The identification of disk fragmentation as a significant influence on GIS performance in selected operations [Dowers et al., 1990]; and

(3) the introduction and use of a wider variety of metrics in characterising and interpreting overall GIS performance (e.g., Disk Fragments, CPU usage, Disk I/O's)

While less systematic and rigorous than the work completed by the Edinburgh researchers, more recent performance comparisons completed by [Hammer et al., 1992] also demonstrate the increased interest in GIS performance assessment in a client-server environment. In particular, this research attempts to quantify the influence of Network File System characteristics on GIS performance. While Hammer et al.'s conclusions in this case may be premature given the limited amount of performance testing apparently carried out, the paper is one of the few to actually demonstrate the link between performance testing results and performance/cost tradeoffs in hardware selection.

This section was designed to present a concise overview of documented and published GIS performance testing efforts undertaken to date. The qualifications are added deliberately, since many other GIS benchmarks were completed during this period as part of normal system acquisition procedures for many different organisations. However, results from most of these are unavailable due to confidentiality constraints. Unfortunately as well, little has been written concerning the testing procedures and approaches employed by consultants in such testing for competitive reasons.

While small, however, the available literature does point to the need for clear objectives and provides a framework for designing defensible approaches to GIS performance testing. Before describing the experiment design in more detail in Chapters 3 and 4, the final section of this chapter briefly describes the original objectives behind this research.

2.6 DEFINITION OF THE GIS/FASTPAC NETWORK RESEARCH

Understanding the strengths, limitations and potential of GIS performance across broadband networks requires an understanding of both the driving visions and technical realities of linking together spatial databases. At a more practical level, it requires an appreciation of the considerations and issues involved in obtaining a reliable indication of performance which would be meaningful to end-users.

This Chapter was designed to provide such an introduction and to give the reader a brief summary of relevant research efforts already undertaken in this field. In particular, the reader should now be able to better understand: (1) the efforts taken in first examining how workers in representative organisations actually use their systems; (2) the factors considered in designing the experiments; and (3) the operational considerations and human factors taken into account when analysing the results.

The final background component to consider is the initial impetus behind the research itself. In June, 1990, the Fast Packet Services Group within Telecom Australia commissioned a series of investigations into the feasibility and the implications of FASTPAC implementation in the GIS/LIS community. The original objectives of this research included:

(1) to characterise the nature of GIS-related operations and data sets common to different user groups in selected organisations dealing with spatial data;

(2) to apply performance testing procedures and network monitoring tools to determine the characteristics of the data traffic resulting from these operations across LANs and MANs (metropolitan area networks) under varying conditions;

(3) to predict the changing characteristics of network traffic load generated by GIS-related operations in different organisations with varying combinations of usage emphasis and network configuration;

(4) to suggest potential effects that higher-speed telecommunications technologies may have on the longer-term information management strategies of an organisation; and

(5) where possible, to model and compare the respective effects of alternative network models on costs and performance levels.

The project was designed and managed by David Coleman from the Centre for Spatial Information Studies (CenSIS) at University of Tasmania under the supervision of Peter Zwart (also of CenSIS). Outside organisations also participating in the research included the Tasmanian Forestry Commission, the

Victoria Department of Conservation and Environment, LANDATA in Victoria and the Sydney Water Board in New South Wales.

Project investigations began in February 1991, with the first series of on-site FASTPAC GIS experiments being carried out from mid-August through late October, 1991. A second series of experiments was authorised during 1992 to examine the influence of specific factors identified during the earlier tests.

Abbreviated summaries of both series of experiments were originally prepared as contract submissions to Telecom Australia ([Coleman, 1992a] and [Coleman, 1992b] respectively). The design and results of both sets of performance-testing experiments will form the basis for Chapters 3, 4 and 5 of this dissertation.

3.0 CHARACTERISING GIS AND NETWORK USAGE IN AN ORGANISATION

Unless an organisation opts for a collection of completely stand-alone systems, the overall nature and volume of network traffic between separate offices will depend on the following institutional and operational factors:

(1) prevailing information systems management strategies within the organisation (implying the logical network configurations in use);

(2) the principal types of activities within the organisation and the respective resources (people, hardware and software) involved in each activity;

(3) geographic locations and mixes of these different resources;

(4) the applications and system operations employed by these groups which generate both routine and peak network traffic loads;

(5) nature and volumes of data being sent among the different user groups as a result of each application.

These characteristics vary between organisations, and no testing strategy can reasonably accommodate all the various permutations involved. However, in order to obtain a realistic indication of GIS performance in a client/server environment across both local and metropolitan area networks, it is important to monitor such performance under a range of conditions and operations. In particular, the operations employed in the performance testing should be representative of those employed in real-world organisations. Since GIS command usage may vary from group to group and from one organisation to another, a defensible approach is required to at least identify "representative" GIS commands for network performance monitoring.

In this research, separate approaches were employed in an attempt to characterise both GIS usage and estimate potential traffic loads within an organisation. To accomplish the former, a form of instrumented GIS usage monitoring using software log files was used to identify the particular GIS commands invoked most frequently and the ones which take up the majority of the user's time. In a separate series of experiments, a network traffic analyzer was employed to quantify overall network loads and traffic patterns under typical operating conditions in an organisation running GIS in a client-server environment.

This chapter describes the design and results of the background research involved in monitoring GIS and network usage in selected organisations, and how this research was used to select the GIS commands used in subsequent testing. In addition to summarising the original investigations into GIS usage and network monitoring, the author discusses some of the limitations and caveats involved in both activities.

3.1 BACKGROUND

3.1.1 GIS User Groups in an Organisation

While the emphasis may change or evolve over time — geoprocessing activities within an organisation can still be generally classified into a finite set of "generic" geoprocessing activities. Allowing for some overlap, GIS software usage within an organisation can generally be classified into the following groups:

Database Loading and Maintenance	Applications Programming
Routine Database Queries	General System Management
Complex Queries and/or Extended Spatial Modelling & Analysis	

To some extent, it may be possible to characterise these classes — or at least the differences between them — by examining how people involved in such activities actually spend their time while using the GIS software. In order to accomplish this, the following factors should be examined, determined and/or quantified:

- the predominant types of operations or commands involved in each generic class of activities;
- the potential amount of network data traffic generated by the principal operations and commands employed by each class;
- any trends which may indicate an existing or potential change in emphasis of the class;
- the number of users in an organisation which fall into each class of activities; and
- the existing or planned location(s) of data holdings within the organisation (centralised, distributed, stand-alone, etc.)...

By identifying and/or quantifying such factors, the relative GIS-related network communication requirements of one group or organisation over another may possibly be predicted in a more reliable manner.

3.1.2 Instrumented Software Usage Testing

While the benefits of having quantitative information concerning system usage may be apparent, obtaining hard numbers can be a problem. Interviews and surveys such as those proposed by Calkins [1991] and others are certainly important, but may often yield misleading results if examined in isolation. Interviewing system users is an essential component of the process, but most users can only be expected to provide detailed information on: (a) work they have done over the past few weeks; (b) a particular set of tasks they repeat periodically; and/or (c) the tasks involved in solving a unique, challenging or otherwise memorable problem. Even in these cases, it is unlikely that the users will recall precise breakdowns of the times involved.

Monitoring the activities of a selected group of users possessing "instrumented" versions of hardware or software has been used in the past to obtain a more quantifiable view of resource usage. During the early 1980's, Telecom Australia conducted an extensive field trial of some new "feature telephones" and logged every keystroke users made on them. From analysing that data, researchers were able to reconstruct user behaviour and facility usage patterns quite accurately. The results of those trials indicated that what users were *actually* doing bore only a very slight relationship to what they reported themselves as doing [Craick et al., 1983]. More recently, instrumented user testing has also been employed to analyse the appropriateness of software structure and to develop and test alternative user interfaces for application software [Bagnara and Rizzo, 1989].

3.1.3 GIS Software Command-Log Records

Many of the larger GIS software packages now have the facility to build records of a person's GIS usage over time. The LOG function in ESRI's *ARC/INFO*[1] software, for example, documents the date and time each command is invoked, the corresponding connect-, CPU- and input/output times required to complete the operation and, in some cases, the name of the file being acted upon [ESRI, 1991]. In *ARC/INFO* , separate command logs are maintained on a "workspace" basis (to record the operations performed by individual users) and on a "coverage" basis (to record the operations performed on each individual map layer). Log entries are typically recorded in ASCII text format and may be easily exported into other text-editing, spreadsheet or database packages.

Software command-log records generated by *ARC/INFO* and other GIS packages can represent an important source of information to system and programme

[1] *ARC/INFO, ARCEDIT, ARCPLOT, ArcView, ARC/COGO, ARC/TIN* and ARC Macro Language (AML) are all registered trademarks representing software products developed by the Environmental Systems Research Institute (ESRI) of Redlands, California, U.S.A.

managers. In addition to helping quantify how geographic information systems are employed individually and collectively in an organisation, they also can indicate how this usage may change over time. However, as will be discussed in subsequent sections, such records usually require careful screening, processing and informed interpretation before they can deliver meaningful information.

A form of instrumented GIS usage monitoring using GIS command log records was used during 1991 to identify representative GIS commands which could be employed in subsequent testing. The implementation of such an approach will be discussed in the next section.

3.2 MONITORING GIS USAGE: EXPERIMENTS

GIS usage files were obtained from fifteen different users in three different organisations employing GIS or facilities management software in their work. Detailed examinations and analyses of system traces and GIS software log files -- supplemented by interviews with staff in each organisation -- were used to identify: (a) the operations most commonly invoked in each user group; (b) the operations which consumed most of the user's time during the sessions; and (c) the operations which generated peak loads on the processor and network respectively.

The following sections introduce the organisations involved in the testing and describe the approach employed in processing and evaluating the log files. Results of early monitoring efforts are also presented, along with a discussion of the strengths and weaknesses of using this approach to quantify GIS usage.

3.2.1 Participating Organisations

Three outside organisations participated in this aspect of the research. These organisations are introduced briefly in the following paragraphs, and reasons for their selection are summarised in Table 3.1.

- The *Tasmanian Forestry Commission* (TFC) represents an organisation which has used GIS since 1984 [Fenn et al., 1991]. Since its state-wide forest inventory is nearing completion, those accessing the GIS include a growing number of routine end-users as well as the more familiar combination of data collectors, planner/analysts and programmers. The Commission was selected in view of its operating a "mature" natural resources GIS database accessible to a large group of users in its Hobart headquarters and (via off-line data transfer) to GIS users in regional offices.

- The GIS Group within the *Victoria Department of Conservation and Environment* employs their GIS for resource inventory and management activities similar to those of the TFC, although they are loading their database over a much longer period [Alexander et al., 1990]. This group tested a number of GIS applications across a high-speed ISDN data communications service during 1990 [Alexander et al., 1992] and have since installed ISDN links between the central facility in Melbourne and regional offices in Kew, Heidelberg and Orbost, Victoria.

- The Regional Information Operation System (RIOS) database in the *Sydney Water Board* now includes property mapping and attribute information -- as well as digitised water and sewer plans -- for a large portion of the metropolitan Sydney area. While a portion of the production staff is still involved with initial database loading and ongoing maintenance, the system has matured to the point that users in 10 different branches of the Board access the database both for routine customer information and for more complex engineering enquiries.

Most such enquiries are remote and must be handled across dedicated communication lines from one of the Board's five other major locations.

Table 3.1: Participating Organisations

Organisation	Reasons for Selection
Tasmanian Forestry Commission (Hobart) *(GIS natural resources application)*	• Database nearing completion with full range of users; • Just beginning to consider alternatives re: data communication to remote offices.
Victoria Department of Conservation & Environment (Melbourne / Kew) *(GIS natural resources application)*	• Prior experience in defining GIS data communication requirements and network testing.
Sydney Water Board (Sydney) *GIS property enquiry and facilities management application*	• Database nearing completion with many routine database users; • Prior experience in defining data communication requirements for a large and varied collection of remote users.

3.2.2 Collection and Summarising of Log Files

To ensure the widest coverage possible, usage information was collected and summarised from 15 different users in the three participating organisations. Although there were some with multiple duties, each of these 15 users fell into one of the five categories mentioned in Section 3.1.1. Depending on the organisation, the log files covered system usage for a period ranging from two weeks (in the case of the Sydney Water Board) to six months (in the case of Victoria Conservation & Environment).

Figure 3.1 contains a typical segment of an unprocessed ARC/INFO workspace log file. All processing for this stage of the research was completed interactively on a microcomputer using commercially-available software. ASCII text versions of the files were initially cleaned and formatted into fields using a word-processing package, and then transferred to a spreadsheet for basic sorting and summarizing.

```
Date    Time    Connect    CPU    I/O    Command

9/5/90    10:16    6    13    8    arcplot

9/5/90    10:18    1    3    0    rotateplot #n-plu100 #n-temp2-r

9/5/90    10:25    3    13    2    HPGL #n-temp2-r AMLC22 1.0
          NOADVANCE NOBANNER XONXOFF 7585 B 0

10/3/90    9:39    3    127    10    buffer gis5d work2 otways 1100
          all  rf7520buf1 rf7520buf4 # # 20 # poly

10/3/90    9:51    12    12    8    ap

10/3/90    10:11    19    440    17    buffer gis5d work2 otways>
          1100 all rf7620buf1 rf7620buf4 # # 20 # poly

10/3/90    10:21    0    1    0    kill fishbuf100

10/3/90    10:37    15    239    18    buffer gis5d work2> otways
          1100 all hydro1100 fishbuf100 bufish
```

Figure 3.1

Example of Unprocessed *ARC/INFO* Command Log File

(GIS Analyst, Victoria Department of Conservation and Environment)

Most log records included 250 to 2000 command records listed chronologically. In the preliminary investigations, the records were re-sorted by command, the time figures summarised, and selected "usage indicators" were then calculated over the entire duration of each log file's history. These indicators included:

(1) Percentage of times a particular command is invoked (in relation to the total number of all invoked commands);

(2) Percentage of CPU time that all the occurrences of one particular command consumes (in relation to total CPU time consumed by that user in that file).

(3) Percentage of Connect time that all the occurrences of one particular command consumes (in relation to the total Connect time consumed by that user in that file).

Standard spreadsheet functions were employed in developing these summaries. Examples of selected summaries are included in Appendix A.

3.2.3 Results

Most prevalent commands can be identified readily from the spreadsheet output. However, bar charts showing the results are useful to illustrate the comparative differences in GIS usage among the various individuals examined. Figures 3.2, 3.3 and 3.4 illustrate the type of comparisons which may be made from this stage of the research.

Since so many *ARC/INFO* commands were employed by the various users overall, the commands were arbitrarily broken down into general categories to make the charts more readable. There is a risk in grouping and displaying all the commands in such a manner; informed users may disagree on any breakdown chosen and the same commands may be used in different contexts depending on the requirements at hand. However, the general trends illustrated by the charts are clear, and the conclusions drawn regarding comparative usage patterns have been confirmed through subsequent interviews with the individuals and organisations involved.

Figure 3.2 summarises the activities of a mapping technician for the Tasmanian Forestry Commission. While many different *ARC/INFO* commands are invoked during the monitoring period, the summary indicates that this person spends most of the time editing graphics data (using the *ARCEDIT* module), editing textual attribute data (using *INFO*), and generating hardcopy output (using the mapsheet formatting and plotting commands within the *ARCPLOT* module).

By comparison, the end-users of the FAUNA database in Victoria Conservation and Environment (shown in Figure 3.3) are concerned primarily with routine database query (using selective data retrieval and display commands in the *ARCPLOT* module).

Finally, Figure 3.4 indicates that the TFC planner employs a wide range of GIS analysis, display and even editing operations in his work during the monitoring period.

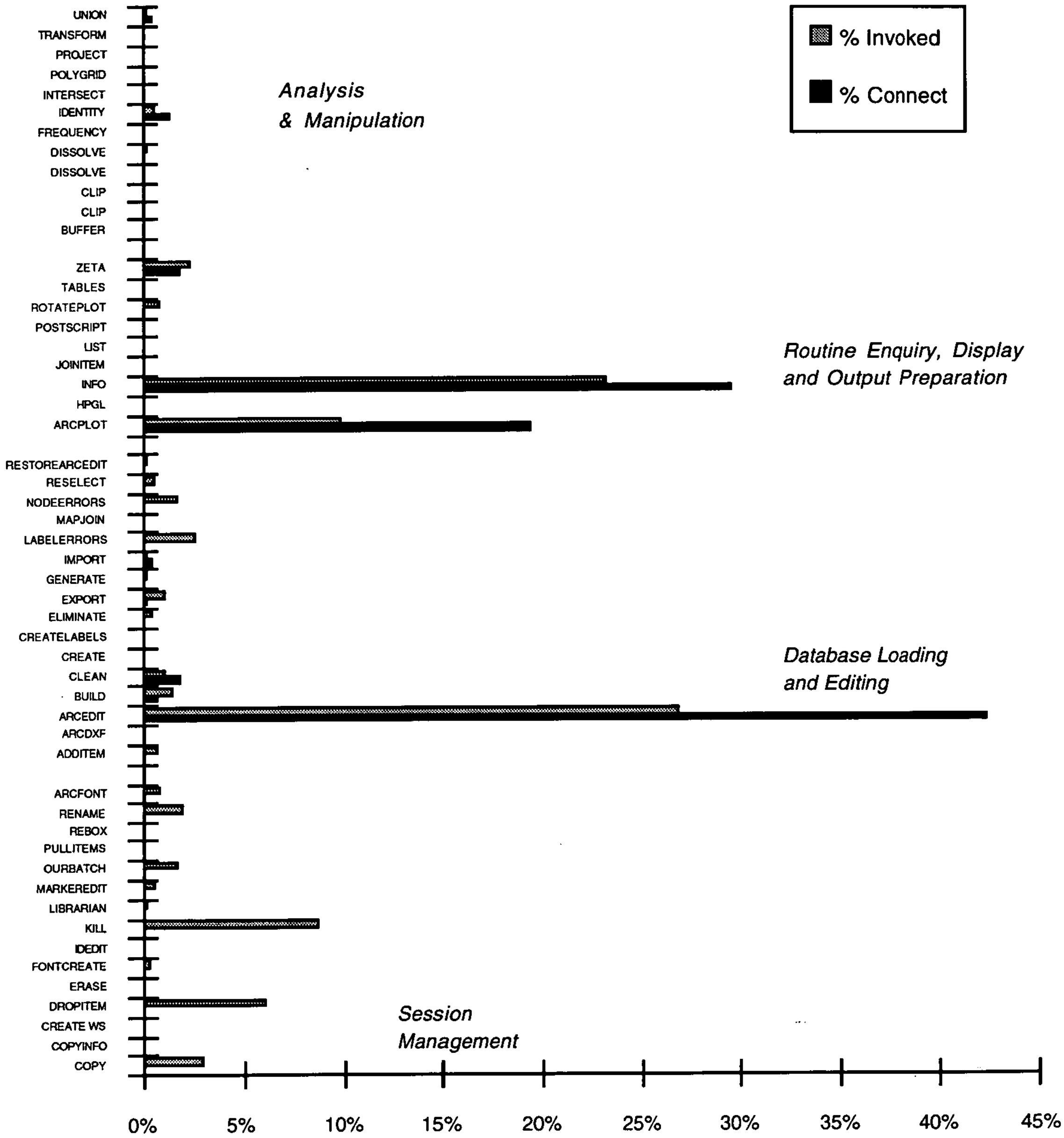


Figure 3.2
Summary of *ARC/INFO* Usage
Mapping Technician, Tasmanian Forestry Commission

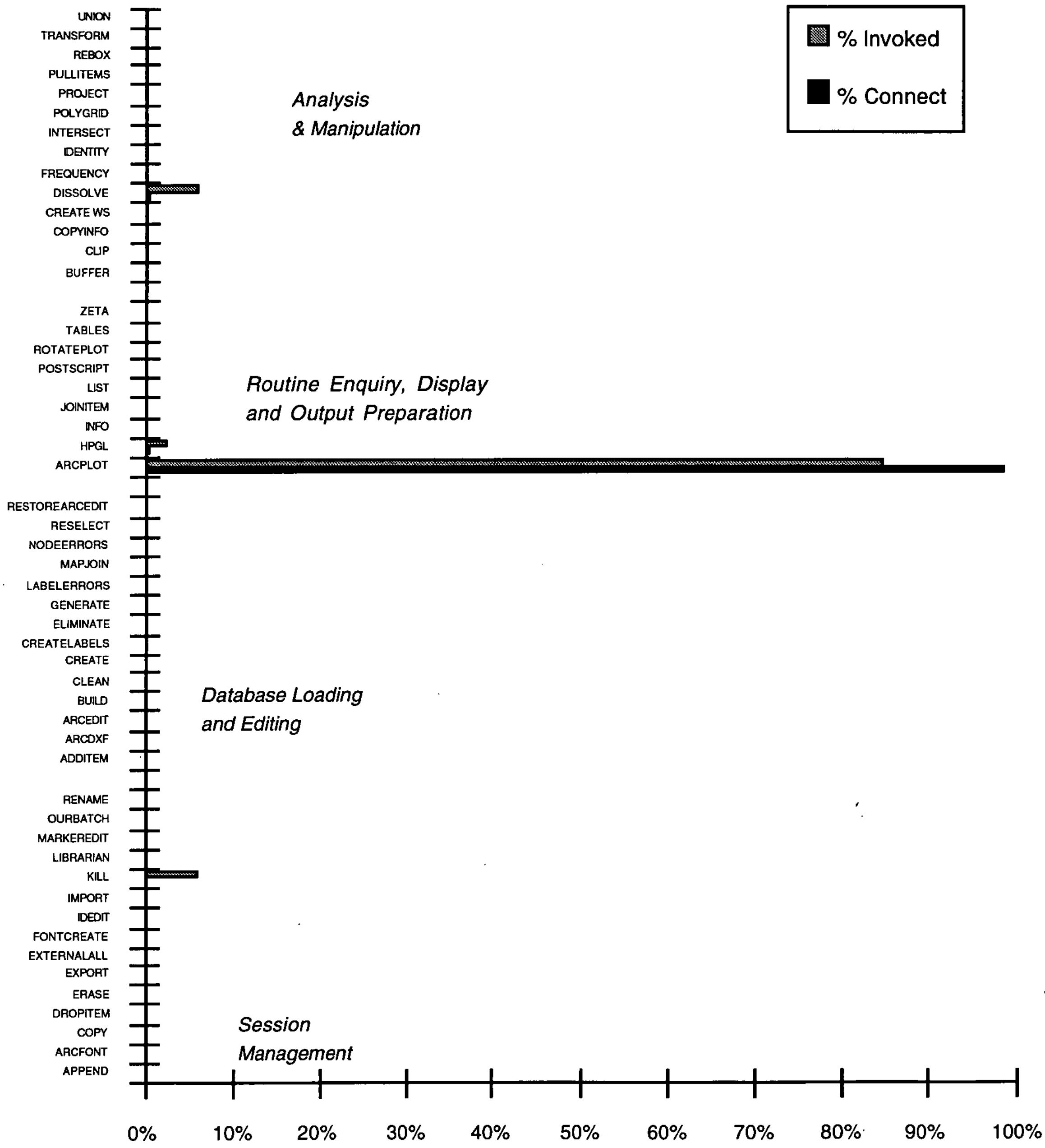


Figure 3.3

Summary of *ARC/INFO* Usage
FAUNA Database, Victoria Dept. of Conservation & Environment

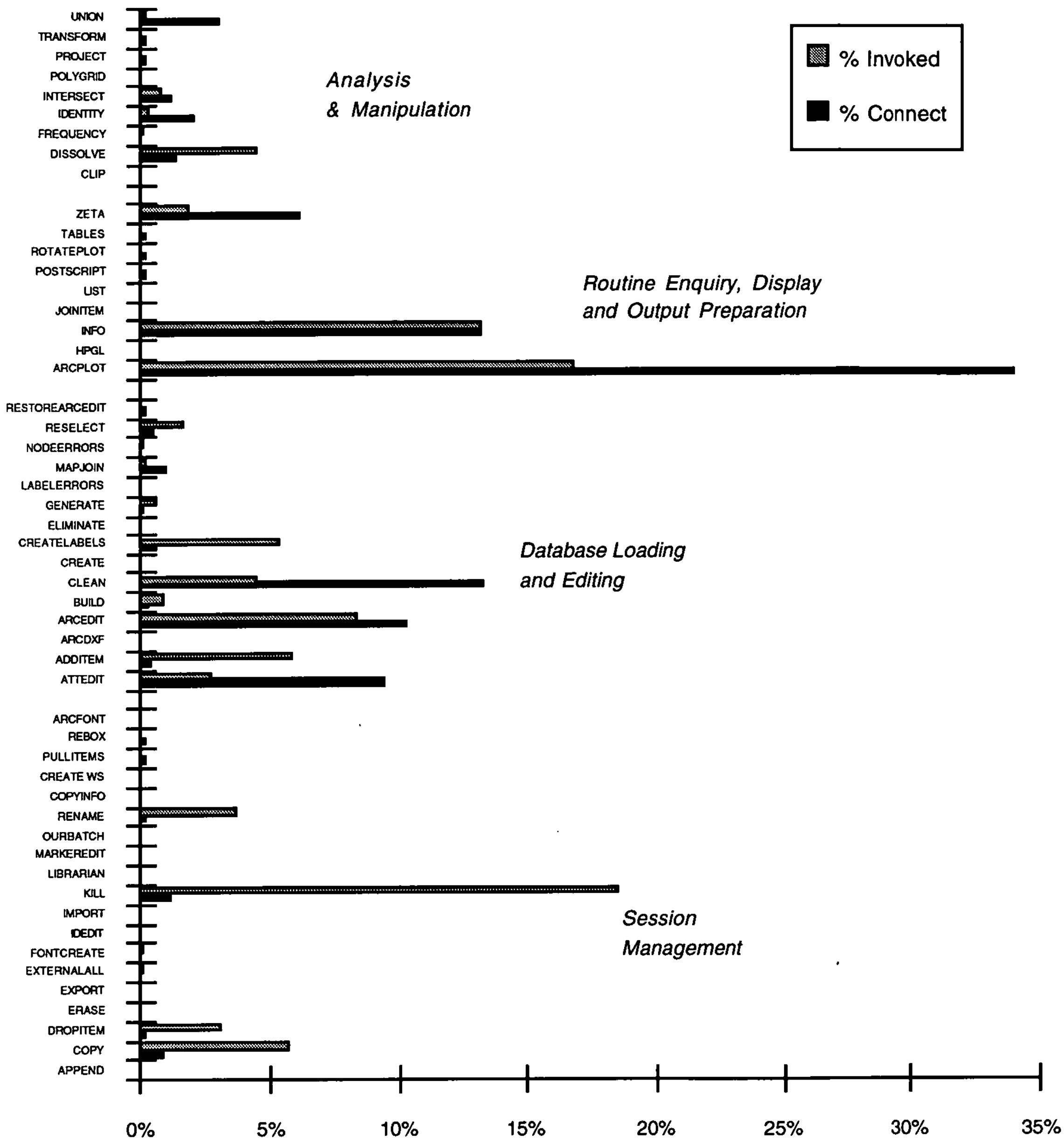


Figure 3.4

Summary of *ARC/INFO* GIS Usage
Forest Planner, Tasmanian Forestry Commission

3.3 MONITORING GIS USAGE: DISCUSSION

3.3.1 Commands Selected for Subsequent Testing

The command-log summaries were useful in helping identify commands most frequently employed for database loading and routine database query, the operations which occupied the most time, and even the most processing- or I/O-intensive operations in each case. For example, it was clear that operations like UNION (polygon overlay), CLEAN (topology construction) and COPY (duplication of *ARC/INFO* coverage directories) were representative of processing-intensive and I/O-intensive operations likely to place a heavy load on the network and/or its components under certain conditions.

It was less apparent *which* editing and display operations should be employed in the testing, since the log files did not record individual *ARCEDIT* or *ARCPLOT* commands. Further discussions with users in the participating organisations failed to reveal any clear indication as to which commands within these modules were used most frequently. Many of the editing operations were either completed with the data stored in local memory or, at worst, placed a minimal load on the network.

Users involved with data collection did agree that the *ARCEDIT* DRAW command was a frequently-invoked display command during editing sessions. Similarly, after polygons with specific attributes were selected using a RESELECT command, it was suggested that end-users employed the POLYGONS and POLYGONSHADES commands in routine query operations using *ARCPLOT*. Finally, while the organisations involved were not using this particular command, the IMAGE command was also selected to determine the comparative times involved in displaying large image files across local and wide area networks.

3.3.2 Limitations of Software-based GIS Usage Monitoring

The log files represented an important source of quantitative usage data and made an important contribution to identifying commonly-employed GIS operations which should be included in subsequent testing. However, experience to date has shown that they cannot be used in isolation. This section discusses software-specific caveats and limitations which must be considered when using this type of information for different applications.

Software-specific factors include any limitations imposed by the *ARC/INFO* software itself. Factors identified during this research included:

- *Log detail and content limitations* — Logging processes, file contents and summarizing procedures will vary depending on the software package in use. Some packages record every command, while others — particularly those which consist of a number of independent but inter-related modules — may only record a portion of the commands used. This is the case in *ARC/INFO*, where individual commands invoked while inside *ARCEDIT, ARCPLOT, ARC/TIN, ARC/COGO* and other modules are not recorded in the log files.

- *Gross errors caused by software or operating system* — Command logging subroutines may have been written early in the life of the software package and may not take into account revisions to operating system capabilities, changes in usage patterns or even new developments in the GIS software itself. In our own experience, bugs in the *ARC/INFO* software prior to Rev. 6.1 caused the logging process to occasionally return vastly inflated and/or negative (-) connect-time figures for certain commands under certain usage conditions. This error was reported and is supposed to have been addressed in Rev. 6.1, but we have not reviewed any log files generated from this version to date. Packages from other vendors may contain their own unique errors, and individual log entries from different users should be examined carefully before any batch processing takes place.

- *User-induced "delays"* — The "Connect Time" figure shown actually indicates the elapsed clock time since the last command invoked. In some cases, it may include substantial amounts of "pause-time" while thinking about a problem, designing a map, talking to someone else, or going off for lunch while remaining logged into the application.

Unless these limitations are addressed in future releases, researchers interested in obtaining more detailed information from *ARC/INFO* users will have to develop a macro operation which: (1) creates or opens a separate log file (an *ARC/INFO* "watch file", for example) when the user launches *ARC/INFO*; (2) records *all* operations completed during the session; and (3) reopens this file (or creates a new one with a unique name) and adds new command records every time *ARC/INFO* is launched.

This would yield an independent and more detailed record of usage. However: (1) the command entries in *ARC/INFO* watch files would still possess some of the timing limitations mentioned above; (2) the resulting watch files would potentially consume a great deal of space; (3) further post-processing and interpretation would be required to turn the watch files into log files; and (4) the development and accounting efforts involved in creating and managing unique watchfiles in a multi-user, client-server environment over a long period may be problematic.

Finally, factors other than software-specific limitations can also affect the value of log-file statistics when attempting to quantify and characterise GIS usage in an organisation. Production processes and priorities often change over time as the organisation's workload evolves. As well, especially in smaller, project-driven organisations, it can be difficult to classify specific individuals with respect to the different categories of GIS usage mentioned earlier [Coleman et al., 1992a].

With all these factors in mind, it is important to have a good understanding of the software, the summarising processes and the respective observation periods employed before summarizing and comparing GIS software usage of individuals and groups within and between organizations. Usage statistics must be interpreted carefully by individuals experienced with the software, job requirements and working conditions, and then supplemented by follow-up interviews with the actual users being monitored.

3.3.3 Summary

An approach to quantifying the GIS usage of selected individuals or groups of individuals in an organisation using existing command logging records and simple-to-construct software tools has been introduced. Initial results confirmed the comparatively wide command of software functions possessed by system managers and analysts, but also suggested that the needs of many end-users are being satisfied by a relatively small number of query and graphics display functions. Most important, by identifying commonly-invoked commands, the process did provide a defensible basis for the subsequent selection of commands which could be employed in the performance testing experiments.

This approach promises to provide a more quantitative estimate of current GIS usage and historical trends than obtainable through interviews alone. Interpreted correctly, such information may yield new insights for front-line production managers and staff, identify training deficiencies, pinpoint bottlenecks in production, and identify most commonly-used files and command combinations. However, in order to be useful, the summaries must be prepared carefully and the analyses interpreted by experienced individuals.

Subsequent research into this subject was carried out at the University of Tasmania through 1992, and a detailed treatment of this work may be found in [Morriss, 1993].

3.4 Monitoring LAN and Inter-Network Traffic

The volume of data transmitted between sites is an important determinant of a customer's wide area networking requirements and ongoing operating costs. However, the process of estimating future levels of WAN data traffic can be problematic and the results unreliable unless some baseline estimates are available which describe the volume and nature of both internal and inter-LAN traffic.

The usage monitoring experiments outlined in Section 3.2 adopt a "bottom-up" approach to this by identifying specific GIS operations and examining the data traffic profiles of each one. While this approach yielded some insight into the behaviour of individual operations, it didn't really capture a meaningful picture of overall network usage.

The "top-down" approach discussed in this section was designed to yield a more complete picture of both network utilisation and inter-network traffic in an operational GIS site. While it is recognised that conditions will vary from site to site, these measurements can offer a general picture of both the average and peak network traffic levels (in terms of bandwidth utilisation) which occur within a workgroup which makes heavy use of its GIS and database resources.

The following sections describe three different sets of network monitoring experiments carried in August, 1992 at two offices of the Victoria Department of Conservation and Environment. After providing some background on the organisation itself and the design of the experiments involved, the results are summarised and the strengths and weaknesses of each approach are discussed.

3.4.1 Site and Network Description

Textual and spatial information processing in the Victoria Department of Conservation and Environment (C&E) is currently evolving from a centralised to a distributed environment. Department-wide databases resided on Prime minicomputers located in C&E offices on Victoria Parade in Melbourne, and database users employed terminals linked to the central minicomputers across asynchronous internal lines and low- to medium-speed data networks.

In late 1991 and 1992, the Department began moving to a client-server environment. A Sun 670 central file server (*Thumper*) was purchased to handle most of the spatial and textual data storage, and asynchronous Tektronix terminals are being replaced by Sun workstations, personal computers and X-terminals linked together and connected to the server across Ethernet LANs. (See Figure 3.6.) Some of the Department's larger regional offices in the Melbourne metropolitan area (including those on Bourke Street and at Kew and Heidelberg) also placed their internal equipment on LANs, while maintaining links into the central file servers using dedicated or dial-up communication services.

A single, dedicated 64 kbit/sec ISDN link connects Kew to Victoria Parade, while a dial-up ISDN line is now being tested between Heidelberg and Victoria Parade. Alexandria and Orbost offices are connected into the central server using lower-speed 9600 bps. lines.

Figure 3.6 illustrates the sub-networks and external connections present at the time the network monitoring experiments were carried out (August, 1992). To optimise ethernet traffic levels in any particular area, users and equipment on the various floors at Victoria Parade are separated onto one of three different sub-nets. Subnet No. 1 contains: (a) the major Sun Server ("Thumper"); (b) a smaller Sun workstation ("LIMS") also used as a server by outside users; and (c) another Sun

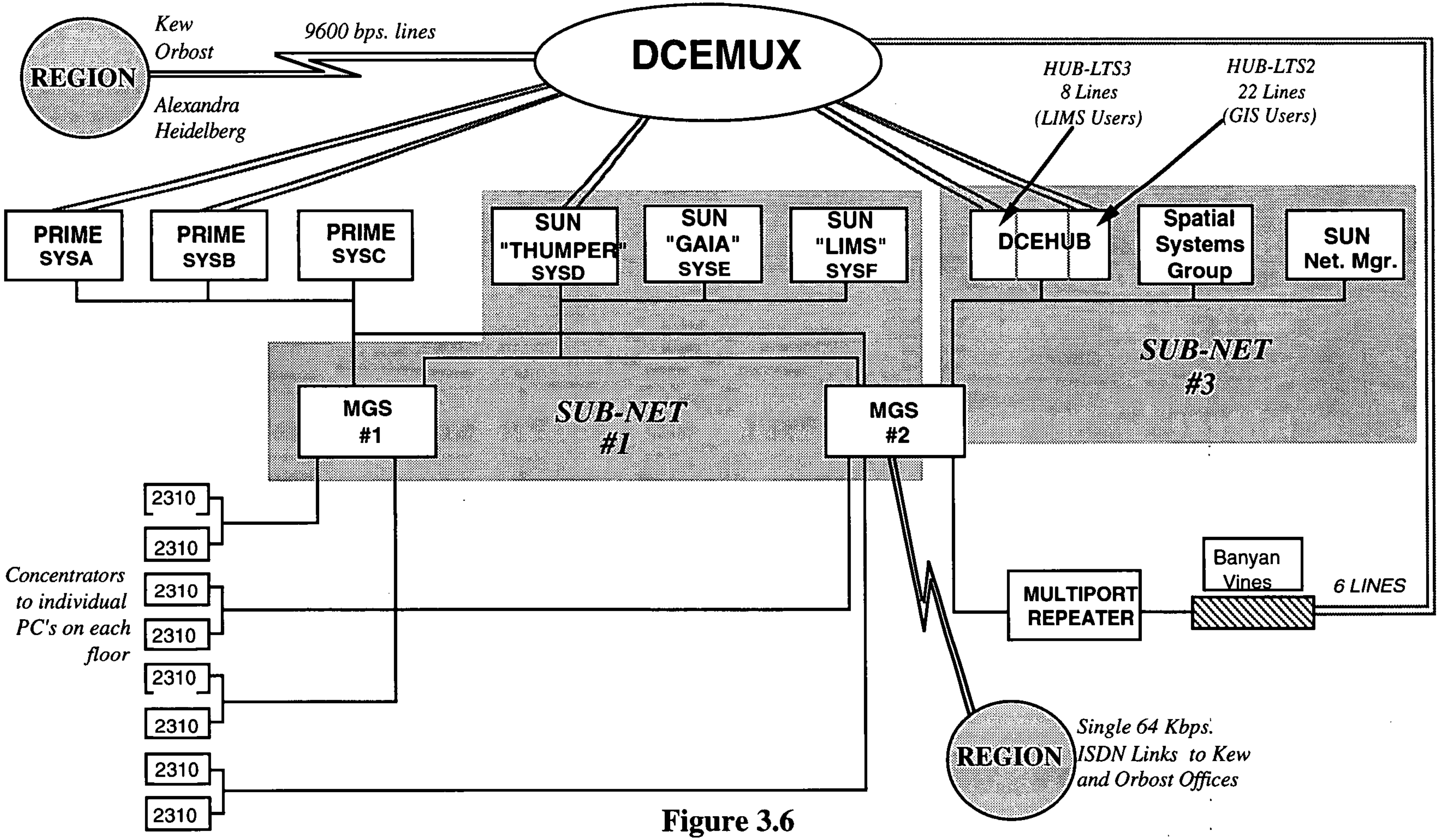


Figure 3.6

Schematic View of Local Area Network
Victoria Department of Conservation and Environment
Victoria Parade, Melbourne

workstation ("GAIA") used as a server for other applications. Subnet No. 3 covers equipment used by the Spatial Systems Group (2 Sun Workstations, two X-terminals, 2 PC's, and two Tektronix graphics terminals) and is linked to Subnet No. 1 across a Cisco router (MGS #2).

Outside users from Kew gain access to "Thumper" and other facilities via an ISDN link across MGS #2 as well. Staff in the Orbost, Alexandra, Heidelberg and other regional offices may access the central facilities via 9600 bps lines connected to the networks via a central multiplexor (DCEMUX). Once inside, these outside users are directed to either the GIS or the LIMS Database across the LAN terminal servers Hub-LTS2 and Hub-LTS3 respectively.

3.4.2 Monitoring Equipment and Tests Employed

During the 5-day monitoring period, a Hewlett Packard HP 18212A LAN Analyzer unit was logically connected to the network at three separate points within the Department. (See Table 3.2.)

Table 3.2: Location of LAN Analyzer Unit for Network Monitoring Experiments
Victoria Department of Conservation & Environment, August, 1992

Location	To Examine:
Subnet No. 1 Victoria Parade	*Overall traffic to and from the Department's two main file servers.*
Subnet No. 3 Victoria Parade	*Traffic generated collectively and individually by the Department's Spatial Systems Group members*
GIS/R.S. LAN Kew Reg. Office	*Internal network load and traffic coming from or going across the ISDN link to Victoria Parade*

In all tests, *HP Performance Analysis Application* software on the Analyzer was used to monitor and summarise specific network- and node-related activity. The parameters monitored and statistics generated fell into three categories, including:

(1) *Network Throughput and Bandwidth Utilisation:* The software monitors and records the number of packets and the corresponding number of kilobytes travelling through the LAN over a given sampling interval. It then accumulates and summarises the amounts over the total measurement period, calculates average values and notes peak values obtained during the test.

In this test, "Percentage Utilisation" over a sample interval is based on the formula:

$$\% \text{ Utilisation} = \frac{\text{LAN Throughput over Sample Interval (Kbits/sec)}}{\text{9,922 Kbits/sec}} \times 100$$

...where 9,922 kbits per second (i.e., ~10 Mbits/sec) is specified as being 100% utilisation of the Ethernet LAN.

Network utilisation was tracked in order to: (a) obtain an idea of network traffic levels vary over the course of a specified interval (e.g., a working day); and (b) obtain an estimate of the overall volume of traffic on the network during the same period.

(2) *Individual Node Traffic Summaries:* The software monitors and records the number of packets (and the corresponding number of kilobytes) transmitted from and received by each node on the LAN over a specified measurement period. These summaries were used to:

(a) identify the most active nodes on the network;

(b) quantify (in kilobytes) the amount of traffic received or transmitted by a given node; and

(c) quantify the amount of traffic received or transmitted across WAN links.

(3) *Node-Network Summaries:* The software monitors traffic between every possible combination of nodes (i.e., "node-pairs") within the LAN, records

the totals for each sampling interval, and then sorts these node-pair combinations by overall traffic volume. Node-Network summaries were used to identify the "high-traffic" segments of the net and — in the case of these experiments — identify which nodes were sending or receiving data across a WAN link.

Each of these tests demanded the full attention of the Analyzer, so data related to all three activities could not be collected simultaneously. Therefore, a combination of short and long tests was used to obtain a more complete picture of network activity in the Department.

3.4.3 Network Utilisation Experiments

Description

The purpose of this series of experiments was to obtain a quick picture of typical and peak load levels on selected LANs. Using output from the LAN Analyzer, variations in LAN utilisation throughout a working day could be observed and a representative range of data volumes flowing across the LAN over a given period could be measured.

Using the basic "Bandwidth Utilisation" functions on the LAN Analyzer, a series of tests recorded the utilisation and throughput levels of Subnets 1 and 3 at Victoria Parade and the GIS/Remote Sensing LAN at Kew. These tests, carried out over the 5-day period, were of varying duration. Basic information concerning test details and overall results may be found in Table 3.3.

Results

While the maximum theoretical transfer rate of an ethernet LAN is approximately 4500 Mbytes/hour (or 10 Mbits/sec), the practical limits usually range from 1200-1500 Mbytes/hour. As indicated in Table 3.3, the observed utilisation levels in all the networks monitored were much lower than these limits. While peak utilisation levels occasionally reached just over 14%, most tests recorded average utilisations of less than 3% during normal working hours.

Table 3.3: Network Utilisation Monitoring Tests

Test Name	Site	Date & Time	Length	Avg./Peak Utilisation	KBytes Transferred	Comments
Prof-01	Victoria Parade	24 Aug. 1135 hrs	18 min.	0.51% Avg. 2.27% Peak	6,622 *(~27,003 Kb/hr Avg.)*	Preliminary check profile for Subnet #1.
Prof-02	Victoria Parade	24 Aug 1355 hrs	60 min.	0.43% Avg. 12.34% Peak	16,830 *(~16,830 Kb/hr Avg.)*	Preliminary check profile for Subnet #3. Only 9 active users.
Net3-01	Victoria Parade	24 Aug 1626 hrs	6 hrs	0.01% Avg. 14.35% Peak	379,700 *(~63,300 Kb/hr Avg.)*	Subnet#3 Profile. File backup activity after 1800 hrs.
Net3-02	Victoria Parade	25 Aug. 0838 hrs	4.5 hrs	2.20% Avg. 11.19% Peak	206,100 *(~45,800 Kb/hr Avg.)*	Subnet#3 Profile. Large component of traffic generated by remote users.
Net1-01	Victoria Parade	26 Aug 0829 hrs	5 hrs	0.21% Avg. 6.00% Peak	148,700 *(~29,700 Kb/hr Avg.)*	Profile of Subnet #1
KNet-01	Kew	28 Aug 0830 hrs	35 min.	0.25% Avg. 14.26% Peak	2,320 *(~4,000 Kb/hr Avg.)*	Early-morning set-up period. LAN usage very low.

The actual volume of data transferred across these LANs (i.e, *Number of Kilobytes Transferred*) ranged from 4 Mbytes per hour when idle up to over 63 Mbytes per hour. The higher values do not place a particularly heavy load on the LANs involved, "spikes" due to heavy utilisation are infrequent, and the peak values observed are still well within the limits required for adequate performance.

Figures 3.7 and 3.8 illustrate utilisation profiles (*Net3-01* and *Net1-01)* across Subnetworks 3 and 1 respectively on successive mornings. In most instances, peak load levels were attributable to intermittent file transfer, backups and graphics display activities. These light loads on Subnets 1 and 3 are due to a recent re-design of the network at Victoria Parade which grouped clusters of users within the building onto four different different subnetworks. According to Department staff, performance delays have been significantly reduced by this modification.

No conclusive evidence concerning cyclic utilisation patterns could be obtained in the limited amount of testing time available. There was some indication that general network usage at Victoria Parade increased during the late morning hours (prior to lunch) and again in late afternoon. However, this phenomenon may have simply been related to activities underway that particular day or week. Obtaining reliable estimates of daily or monthly cycles in network utilisation (if they exist) would require a much longer and continuous observation period.

At Kew, most of the peaks observed were caused by file-transfer activities rather than by graphics displays. Given the amount of graphics editing underway in this office, terminal traffic is characteristically light until a graphic or image must be redrawn on the screen. As a result, average network utilisation stayed below 1% in most cases measured, with occasional spikes of 12-15% during file transfers.

These findings suggest that — while significant peaks do occur during file transfers and back-up activities — the *sustained* levels of data traffic across these particular LANs remain very light. The presence of low average traffic levels across local area networks running text-based applications has been well-documented by [Shoch et al., 1980] and [Teorey et al., 1990], among others. While the results of these experiments imply the same may also hold also true in a GIS environment, other researchers suggest that average LAN utilisation in mixed application environments may actually be much higher than the levels encountered here ([Healey, 1994]; [Fowler et al., 1991]).

On a more immediate level, the average and peak utilisation values observed here should be kept in mind when GIS performance is examined under different network traffic conditions (to be discussed in Section 5.3.).

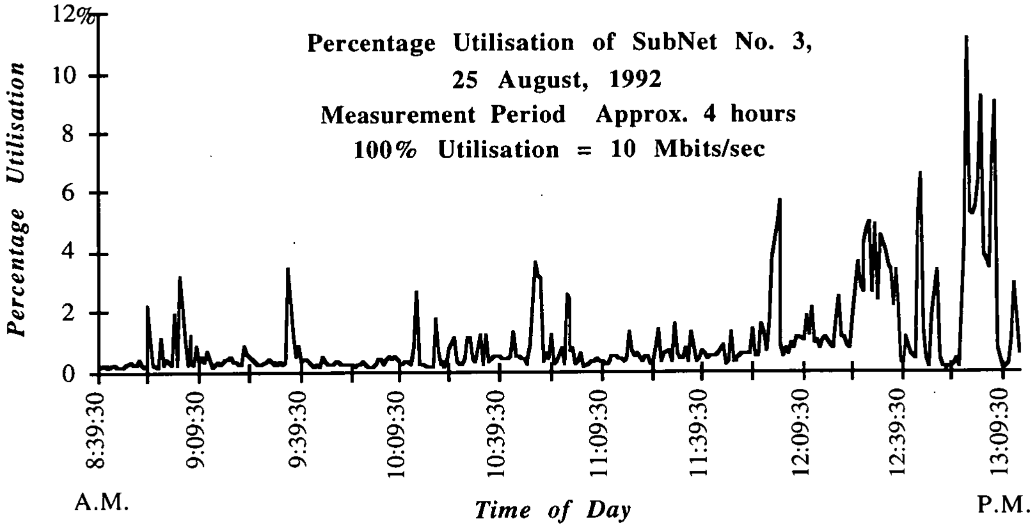


Figure 3.7

Utilisation Profile: Subnet No. 3, 250 Victoria Parade

Victoria Department of Conservation & Environment

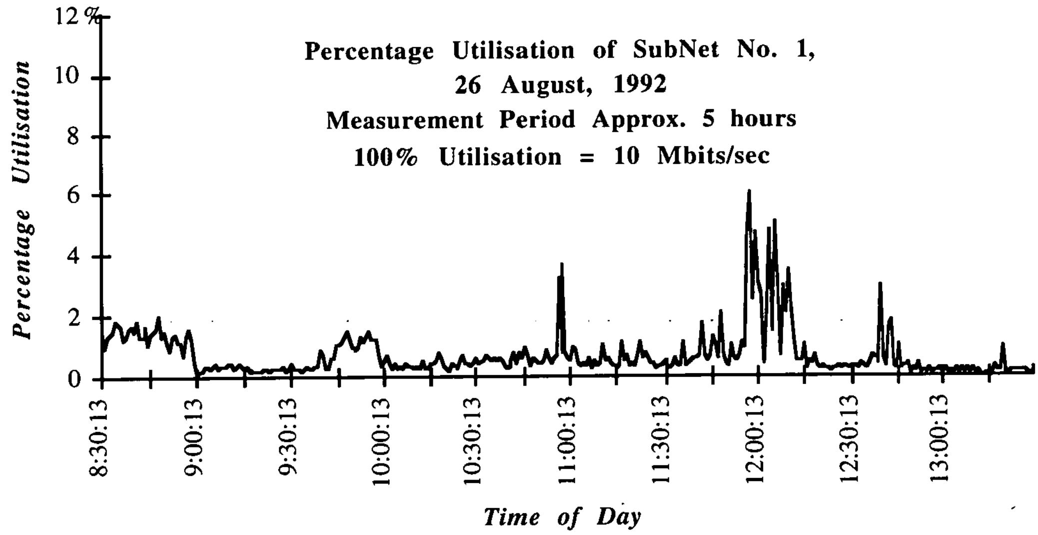


Figure 3.8

Utilisation Profile: Subnet No. 1, 250 Victoria Parade

Victoria Department of Conservation & Environment

3.4.4 Results: Inter-Network Traffic Monitoring

Purpose

The purpose of this test was to determine if we could use LAN Analysis software to gain a clearer picture of how extensively the wide-area network links were being used in the Department. Two sets of experiments were employed to accomplish this:

(1) Measuring GIS-related traffic from low-speed lines coming into Victoria Parade from Department regional offices across two LAN Terminal Servers; and

(2) Measuring GIS-related traffic across an ISDN link between Victoria Parade and the Department's regional office in Kew.

Lower-speed Services into Victoria Parade

GIS users at Department regional offices on Bourke Street and in Orbost, Heidelberg and Alexandra, among others, all tie in to *ARC/INFO* software and data which resides on the central Sun server (*Thumper*). While higher-speed ISDN links are either now being installed or being planned, all these offices currently use asynchronous graphics terminals tied to the server across lower-speed 9600 bps. communication lines. Depending on the database being accessed (LIMS or GIS), these remote users enter Subnet #1 at Victoria Parade across one of two LAN terminal servers HUB-LTS2 or HUB-LTS3. (See Figure 3.6.)

Kew — Victoria Parade ISDN Link

By design, copies of the *ARC/INFO* GIS software and selected portions of the GIS database now reside at both Victoria Parade and the Kew regional office. However, staff at the Kew office still log into the central server in Victoria Parade when they wish to: (a) access data holdings not stored locally; (b) reduce the load on the local server; (c) perform routine backup and database maintenance duties;

or (d) assist GIS users in other offices who may only have access to the central server. In those cases, the central server is accessed via the 64 Kbps. ISDN link.

Description

Two sets of experiments were performed over a 4-day period. During the first two days, packet traffic to and from the LAN terminal servers on Subnet No. 3 was observed and measured at different times using the "Node-Network Summary" software on the LAN Analyzer. (See Table 3.4.) For the final two days, packet traffic between the ISDN router KEWHUB and all other nodes on the LAN at Kew was observed and summarised using the same software. (See Table 3.5.)

Table 3.4: Inter-Network Traffic Monitoring Tests
(Between Victoria Parade and Regional LIMS and GIS Users)

Test Name	Site	Date & Time	Length	KBytes Into Vic. Pde	KBytes Out from Vic. Pde	% of LAN Traffic	Comments
Node-01	Subnet 3, Victoria Parade	24 Aug. 1511 hrs	30 min.	1,155 *(~2,310 KByte/hr Avg.)*	1,690 *(~3,380 KByte/hr Avg.)*	34%	High component of active remote users, with two major local users inactive for most of test.
Node-03	Subnet 3, Victoria Parade	25 Aug 1400 hrs	151 min.	9,494 *(~3,770 KByte/hr Avg.)*	10,046 *(~3,990 KByte/hr Avg.)*	4%	Afternoon -- Major activity from single local user, although many remote users logged in.
Node-04	Subnet 3, Victoria Parade	25 Aug 1700 hrs	360 min.	2,170 *(~360 KByte/hr Avg.)*	2,492 *(~420 KByte/hr Avg.)*	6%	Evening -- Low activity after normal working hours.

Table 3.5: Inter-LAN Traffic Monitoring Tests
(Between Kew Regional Office and Victoria Parade)

Test Name	Site	Date & Time	Length	KBytes into Kew from Vic. Pde	KBytes from Kew to Vic Pde.	% of LAN Traffic	Comments
KNode-01	Kew	26 Aug. 1700 hrs	12 hrs.	2,213 *(~180 Kb/hr Avg.)*	885 *(~80 Kb/hr Avg.)*	1%	Evening -- Very low activity after normal working hours.
KNode-02 and KConn-01	Kew	27 Aug. 0930 hrs	98 min.	40,112 *(~24,560 Kb/hr Avg.)*	2,400 *(~1,470 Kb/hr Avg.)*	30%	Morning -- Very Heavy GIS-related file transfer activity from Server at Vic Pde.
KConn-02	Kew	27 Aug 1115 hrs	60 min.	3,754 *(~3,750 Kb/hr Avg.)*	5,047 *(~5,050 Kb/hr Avg.)*	20%	Later same morning -- File transfer activity finished, but three users logged into Vic Pde. for GIS editing activities.
KConn-04	Kew	27 Aug. 1336 hrs	10 min.	16 *(~100 Kb/hr Avg.)*	0 *(0 Kb/hr Avg.)*	2%	Lunch Period -- No staff activity across to Vic. Pde. at all.
KConn-05	Kew	27 Aug 1354 hrs	18 hrs.	12,207 *(~680 Kb/hr Avg.)*	6,387 *(~350 Kb/hr Avg.)*	3%	18-hour period covering entire afternoon and evening shifts
KConn-06	Kew	28 Aug 0923 hrs	6 min.	776 *(~7,760 Kb/hr Avg.)*	877 *(~8,770 Kb/hr Avg.)*	71%	Morning -- Mostly terminal traffic into Vic. Pde..

Results

In the tests at Victoria Parade, the normalised packet traffic between the central servers and the various regional offices (connected through LAN terminal servers HUBLTS2 and HUB-LTS3) ranged from 360 Kbytes to 3,770 KBytes per hour. Given the small "entry pipe" provided by the LAN terminal servers (9600 bps), one should expect that the traffic coming in from the regional offices should usually represent only a small component of the overall traffic on the network. The figures bear this out: during the observation periods, inter-network traffic represented less than 10% of the overall total in two of the three tests, with a higher percentage (34%) occurring only when two of the major local users were off the system.

Usage across the ISDN link between Kew and Victoria Parade (via the router KEWHUB) was somewhat higher, but much of its capacity was largely underutilised during most of the testing period. Normalised packet traffic rates ranged from virtually zero during an idle lunch break to over 24 Mbytes per hour during a series of large file transfers. The relative proportion of traffic going across the ISDN link varied substantially over the six tests, ranging from 1% (after normal working hours) to 71% (at the beginning of working hours).

3.4.5 Discussion

In the time available for the experiments, it was only possible to determine which tests were more valuable than others for traffic projection purposes. Of the various tests employed in these experiments, it appears that the *Network Throughput & Bandwidth Utilisation* monitoring and the *Node/Network Summaries* delivered by the LAN Analyzer (described in Section 3.1.2) were the most useful in quantifying important characteristics of LAN and inter-LAN traffic at the sites visited. The former provided a time-stamped "snapshots" of internal LAN utilisation and indicated the overall amount of data traffic carried over a given period of time. The latter identified the most active nodes on the LAN, indicated "heavy-traffic" pairs along the network and quantified the amount of inter-LAN traffic.

The following caveats should be kept in mind when considering the above results:

(1) The results of this testing provide a *snapshot* of LAN and Inter-LAN traffic at the two sites *over a very short period only.* Given the influences of staff vacations and workload variation, a much longer and more intensive period of

monitoring should take place in order to obtain statistically-reliable estimates of representative traffic loads and utilisation cycles in this Department. In fact, discussions with systems support staff suggest that the traffic levels measured in these tests may be lower than "normal" for the networks in question given a smaller-than-normal number of users for part of the monitoring period.

(2) GIS Traffic only accounts for a portion (albeit significant) of the overall traffic loads on the various networks and sub-networks observed;

(3) Because the network at Victoria Parade had been subdivided to improve performance, the Analyzer had to be alternately placed on different networks at different times. Therefore, an overall picture had to be compiled from these separate, non-concurrent measurements.

(4) Since traffic from regional offices all came in through one of two LAN terminal servers, specific regional users could not be identified through the LAN Analyzer. While individual system users could be identified at a given time by system logs, the Analyzer output indiscriminately aggregated all traffic coming in or going out across these terminal servers. Therefore, relative traffic loads to and from different regional offices could not be determined.

(5) Network-node summaries represent aggregate data volumes transmitted or received over the course of the entire sampling period. The addition or absence of one or two large file transfers or GIS display operations in any particular session, for example, could significantly affect these summaries and — by implication — the figures shown in Tables 3.3, 3.4 and 3.5.

As mentioned earlier, the Analyzer was capable of running only one type of test at a time. Ideally, *two* Analyzers should be used to run the above tests concurrently in order to get a more complete picture of network/user activity. Observers could then relate particular "spikes" in traffic to activity between particular nodes by comparing the time-stamped output from one Analyzer with that of the other. Whether or not this degree of rigor is required or not would depend on the potential value of customer or contract, the complexity of the site or organisation being monitored and the willingness of the potential customer to allow such on-site monitoring.

Finally, any results obtained and conclusions drawn from such measurements can only represent levels and patterns of *existing* traffic. As such, they can only be considered as a baseline for estimated levels which will occur when a higher-speed network is in place. In this particular case, for example, traffic levels observed reflected the practices of remote terminal users tied to a central system by relatively low-speed lines. Given higher-speed lines and an extended client-server architecture, one can assume that usage patterns in regional offices would change (depending on whether stand-alone workstations or X-terminals were purchased) *and* that overall network traffic levels to and from the central server would increase.

3.5 CONTRIBUTION OF GIS AND NETWORK USAGE OBSERVATIONS TO OVERALL RESEARCH OBJECTIVES

As discussed in the preamble, the two approaches outlined in this chapter represent "bottom-up" and "top-down" approaches to examining GIS and network usage completed in 1991 and 1992 respectively. This final section places these experiments once again into a larger context, and then summarises the lessons learned and the contributions of these experiments to the overall research objectives.

The original purpose of 1991 GIS usage experiments was to identify a selection of GIS operations (in fact, *ARC/INFO* commands) which could defensibly regarded as being "representative" of operations carried out in actual GIS production sites. The results (as discussed in Section 3.3.3) did provide significant insight into GIS usage characteristics and trends in different organisations. However, it is problematic to combine such statistics from a group of users in order to make anything more than general predictions concerning the resulting server and network loading in an operation. While it would be possible to develop simulated workloads given the log file data available, the effort involved in producing a realistic simulation would take far longer than simply taking actual measurements of the application itself under multi-user conditions.

By the same token, the network monitoring measurements described in Section 3.4 were originally designed to quantify and characterise network loading in a "typical" GIS production environment, with a view to creating "network background traffic load" profiles which could be used in subsequent performance testing experiments. Scheduling conflicts prevented these measurements from taking place until after the FASTPAC testing was completed, and other criteria were eventually employed to simulate varying traffic loads. However, even if the profiling was carried out in time, there turned out to be so much variation in the individual traffic profiles measured that — while instructive — only general conclusions could be drawn.

Practical network capacity planning does not require such detailed information to proceed. Rather, network planners require a more general picture of usage consisting of, e.g.: (a) projected levels of data to be transmitted across a network on daily, monthly or weekly basis; (b) the times, levels and estimated duration of "spikes" in data traffic through the day; and (c) the maximum allowable response time for specific operations. At present, such information is either estimated using information gathered through interviews and/or calculated using rough data transmission figures.

While neither set of experiments fulfilled all the objectives originally laid out, they nevertheless did provide insights that make an important contribution to both the overall objectives of this research. Specifically:

- Some of the more commonly-invoked commands identified from the instrumented GIS usage monitoring were employed in subsequent FASTPAC/GIS performance testing (to be described in Chapters 4 and 5).

- The approach and summation techniques developed for instrumented GIS usage monitoring experiments have been refined and employed in wider investigations of GIS usage in Australia and Canada [Morriss, 1993]. Using such techniques to identify and quantify trends in GIS usage in an organisation holds promise in the areas of procedural optimisation, determination of employee training requirements and the development of customised systems.

- The network monitoring experiments provided an independent assessment of the fluctuations in network utilisation which may be present over an operating LAN through the course of routine operations. While specific levels and ranges will vary depending on the organisation, these order-of-magnitude figures would prove useful in later experiments which examined how GIS performance diminished under increasing network traffic loads.

- Beyond just the spatial data handling context, the network monitoring approach described in Section 3.4 offers a defensible approach to quantifying current overall network usage. In cases where (for example) an organisation is considering upgrading its data communications services, output from the LAN Analyzer can be used to corroborate or refute information gathered through interviews and rough calculations of existing load levels.

  Ideally, further research would be required to refine this approach and test it under varying conditions in different organisations. However, while potentially of interest in a network management context, such work falls outside the scope of this research.

The investigations described in this Chapter provided necessary insight into both GIS software usage and overall network utilisation in a client/server environment where GIS was the primary application. It is now necessary to take these inputs and incorporate them into a series of tests which rigorously examine GIS performance under carefully controlled conditions. The next chapter introduces these tests by discussing their design, the equipment and data sets employed, and the manner in which they were ultimately carried out.

4.0 PERFORMANCE MEASUREMENT: EXPERIMENT DESIGN

Particularly over the past five years, rigorous GIS performance evaluation has become an important area of research. However, while the implementation of geographic information systems software in a client-server, local area network environment is now commonplace, most GIS performance-testing research efforts to date have been based on *stand-alone* workstations and PC-based systems.

As technology now moves beyond high-speed LANs towards broadband *wide* area networks, organisations are already considering the use of such technology to connect widely-scattered users with corporate datasets, software and computing resources. Accordingly, while there is much that can still be done to "map out" GIS performance space on stand-alone systems, research efforts must begin to evaluate the same performance in a client-server environment on both local and wide area networks.

The previous chapter identified representative GIS operations routinely performed in an organisation and quantified the nature of GIS-related background traffic which may be encountered within and between different local area networks. Using this information, rigorous performance testing can then be applied to determine a reliable estimate of the respective times required to complete GIS and Unix operations on: (1) a stand-alone workstation; (2) a client-server configuration across a local area network; and (3) a client-server configuration linking two LANs via the FASTPAC high-speed metropolitan area networking service.

This chapter discusses the manner in which this performance testing was designed and carried out. In addition to describing the equipment employed, it summarises the characteristics of the representative data sets, the hardware and software, the wide-area network connections and the procedures employed in the experiments.

4.1 RESEARCH OBJECTIVES

The original hypothesis leading to this research was that broadband communications networks would provide the performance necessary for an organisation to satisfactorily support the GIS application and data management requirements of a geographically-dispersed organisation from a single location. With this hypothesis in mind, the specific research questions to be addressed here then become:

(1) How can the performance of specific GIS (and related) applications be evaluated in a client-server environment across a broadband network?; and

(2) What constitutes "satisfactory performance" in such an environment?

With these questions in mind, the major goal of this stage of the research is to design and implement a methodology which will indicate to GIS end-users whether or not a specified networking service can provide satisfactory response-time performance for a given set of operations. Further, it is important that any evaluation methodology be conducted — and the results expressed — in a manner which will be relevant and understandable to end-users.

Assuming that:

(1) the average times required to complete such operations on a stand-alone workstation (t_{ws}) or across a well-behaved local area network (t_{lan}) are currently considered to be acceptable to most users; and

(2) there exists some agreed-upon delay factor (Δ_D) that, if exceeded, will result in subsequent GIS performance being deemed unacceptable by the same users...

THEN the relative increases in time (Δ_T) involved in executing the same operations across a *broadband* network (t_{bbn}) should provide a reasonable indication of whether those particular operations can be satisfactorily carried out in such an environment.

Expressed in another manner, Δ_T must be equal to or less than Δ_D, where...

$$\Delta_T = \frac{(t_{bbn} - t_{lan})}{t_{lan}} \times 100$$

Based on these assumptions, it is proposed that experiments be conducted under controlled conditions which would measure the corresponding times taken to perform the same operation on a stand-alone workstation, across a LAN, or between LANs across FASTPAC. The *absolute and relative differences* in execution times – measured using a combination of performance-timing utilities and empirical calculations – would serve as the basis for comparing the different approaches.

Since they will vary widely depending on the hardware and software in use, it is important to make the distinction between using *relative and absolute time differences* as opposed to *individual execution times* in these experiments. Execution times alone are usually not meaningful as general indicators of performance, since many factors can combine to influence execution times and file transfer speed across communication networks ([Boggs et al., 1988], [Sloan et al., 1992] and others). As discussed in Section 2.6.2, these factors may be grouped into different categories, including the network and workstation hardware employed, the operating system and communications protocols in use, the data structures and techniques employed by the application software, and the nature of other loading on the networks involved.

Provided these effects of these factors can be isolated or controlled during testing, however, monitoring and comparing corresponding execution times can indicate the relative magnitude of GIS performance *differences* between various stand-alone, LAN and WAN configurations. In certain cases — if enough operations are monitored in different configurations — comparing sets of differences may also help identify the relative magnitude of the factors affecting performance.

The majority of GIS users in most installations have an application-specific rather than computer science background. With this in mind, it has been assumed that most end-users will be more concerned with physical and logistical factors over which they have some degree of control than those dealing with optimisation of hardware, application-to-operating system linkages or communication protocols. In these particular tests, then, only the nature of the client workstations, the data volumes, server activity and background network traffic conditions are varied. No assumptions are made concerning more complex factors which may affect network file transfer rates (e.g., queuing delays or processing overheads).

Given the context of the research and the hypothesis being tested, it is believed that taking such an external "black-box" approach to GIS performance testing is valid for this particular case. Support for such an approach to GIS evaluation may also be found in recent research undertaken by Wagner [1991], Healey et al. [1991], Hawke [1991a] and, to a some extent, by Hammer et al. [1992].

Based on this initial concept, two sets of performance trials were subsequently carried out through both 1991 and 1992. The following sections describe the equipment, data, configurations and operations actually employed and discuss the potential limitations of this approach as originally seen.

4.2 EQUIPMENT AND NETWORK CONFIGURATIONS

All tests were carried out under controlled conditions using identically-configured client and server workstations and consistent versions of the operating system and GIS application software. Dedicated LANs and FASTPAC connections were employed to eliminate the influence of any unwanted external traffic; levels of any additional network traffic loads were varied precisely using a LAN traffic generator program.

This section describes the equipment used in the FASTPAC/GIS performance testing. After outlining the hardware and software employed, the alternative arrangements of equipment within the local- and metropolitan area network configurations are described.

4.2.1 Hardware and Software

In the first series of tests (completed in September/October 1991), identical Sun IPC workstations were used in order to minimise any bias due to different hardware components. The Sun IPC server *A* was configured to: (a) provide both the UNIX and NFS operating system software to the two client workstations *B* and *C* across the FASTPAC cloud; (b) act as the single host for the GIS software; and (c) provide allocated "scratch" space for the diskless client *C*. Characteristics of each of the workstations involved are summarized in Table 4.1.

The 1992 series of tests was carried out in Melbourne in June. Identical DECstation 5000 workstations were provided by Telecom Australia, and a Labtam CT200 X-Terminal was used to investigate comparative traffic generated by such equipment. This time, the UNIX, NFS and X-terminal driver software was present on both Server *(A)* and the Client *(B)* workstations. X-Windows manager software for the X-terminal *(C)* resided on both *A* and *B*, and could be downloaded from either host at the beginning of each terminal session. Characteristics of each unit in the 1992 tests are summarized in Table 4.2.

Table 4.1: Summary of Workstation Equipment Employed in 1991 Experiments

A	B	C
Sun IPC Workstation *12 Mb. Main Memory* *Default Sun NFS settings* *(BIOD-8; NFSD-8)*	**Sun IPC Workstation** *12 Mb. Main Memory* *Default Sun NFS settings* *(BIOD-8; NFSD-8)*	**Sun IPC Workstation** *12 Mb. Main Memory* *Default Sun NFS settings* *(BIOD-8; NFSD-8)*
207 Mb. Disk Drive *16 msec Average Seek Rate* *4.0 Mb/sec Data Transfer Rate*	**207 Mb. Disk Drive** *16 msec Average Seek Rate* *4.0 Mb/sec Data Transfer Rate*	Reconfigured as Diskless Client to Server A
669 Mb. Disk Drive *16 msec Average Seek Rate* *1.8 Mb/sec Data Rate*	Reconfigured as Diskfull Client to Server A; Only a portion of Operating System (O/S) resides here.	Dependent on A for all O/S and application operations; Portion of Disk on A partitioned for use by C.
SunOS UNIX		
***ARC/INFO* Rev 5.01 GIS Software**		

Table 4.2: Summary of Workstation Equipment Employed in 1992 Experiments

A	B	C
DECstation 5000 Workstation	**DECstation 5000 Workstation**	**Labtam CT200 X-Terminal**
24 Mb. Main Memory *Two Sets of NFS settings* *(BIOD-8; NFSD-0)* *(BIOD-8; NFSD-8)*	*24 Mb. Main Memory* *Two Sets of NFS settings* *(BIOD-8; NFSD-0)* *(BIOD-8; NFSD-8)*	*8 Mb. Main Memory;* *2 Mb. Video memory*
Two 665 Mb. RZ56 Disks *16 msec Average Seek Rate* *1.875 Mb/sec Data Rate*	**One 665 Mb. RZ56 Disk** *16 msec Average Seek Rate* *1.875 Mb/sec Data Rate*	
DEC ULTRIX	**DEC ULTRIX**	
***ARC/INFO* Rev 5.01 GIS Software**		

In both sets of tests, the GIS application software *ARC/INFO*™ Rev. 5.01 was present *only* on Server *A*.

Finally, a Hewlett Packard 18212A LAN Analyzer unit was also included in both sets of testing. This unit is a stand-alone microcomputer. Connected to the same LAN as one or both of the workstations, it runs custom software which gives the operator the capability to monitor overall LAN throughput and utilisation, traffic to and from specified nodes on the network, or traffic between two specified nodes on the network [Hewlett-Packard, 1989]. Another portion of the software allows the unit to act as a "network traffic generator", where the user can specify the nature and degree of background traffic between two nodes within the same LAN or on connected LANs.

™ *Arc/Info* is a product of the Environmental Systems Research Institute (ESRI), Redlands, California, U.S.A.

4.2.2 Network Connections

1991 Experiments

The local area network configuration was used as a baseline for user expectations within a client-server environment. (See Figure 4.1.) One wide-area network configuration ("FP[A]" — See Figure 4.2) connected two LANs within the same FASTPAC sub-network to simulate connecting two LANs in different buildings but both residing in the central business district. A second WAN configuration ("FP[B]" — See Figure 4.3) simulated connecting two LANs 25-30 km. apart (e.g., across a metropolitan area). In both cases, the allocated bandwidth of the FASTPAC service was 10 Mbits/sec.

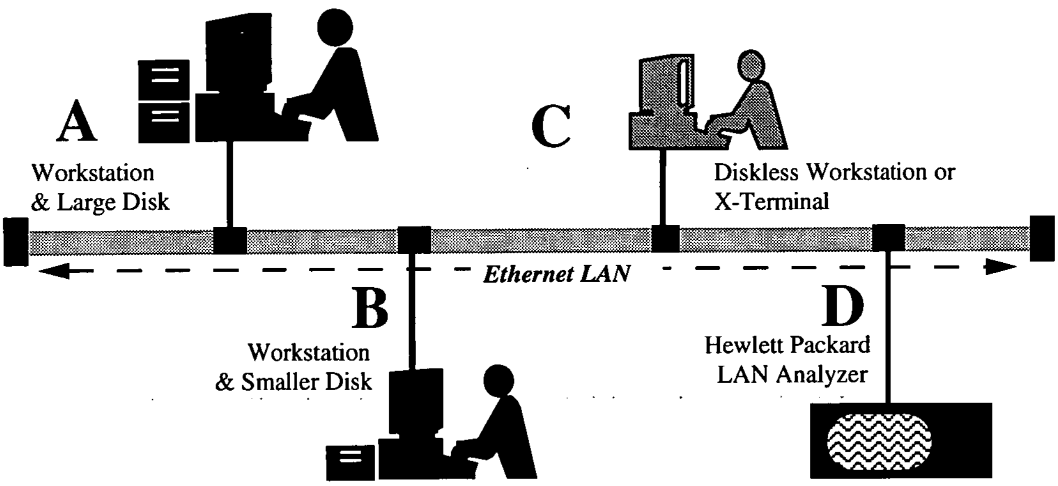


Figure 4.1
LAN Testing Configuration

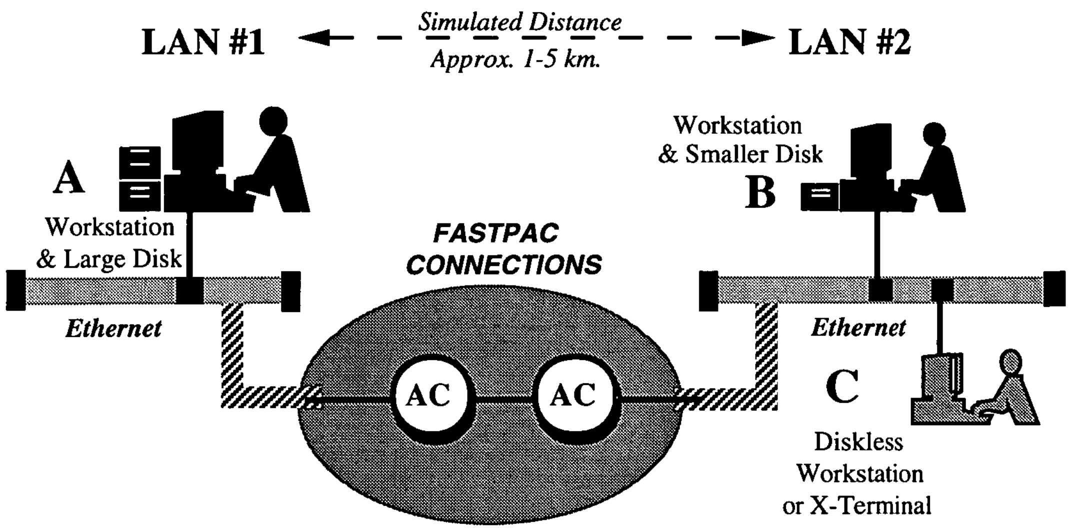
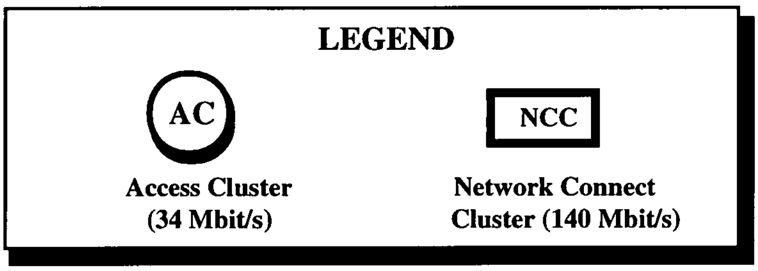


Figure 4.2: FASTPAC *FP[A]* Configuration



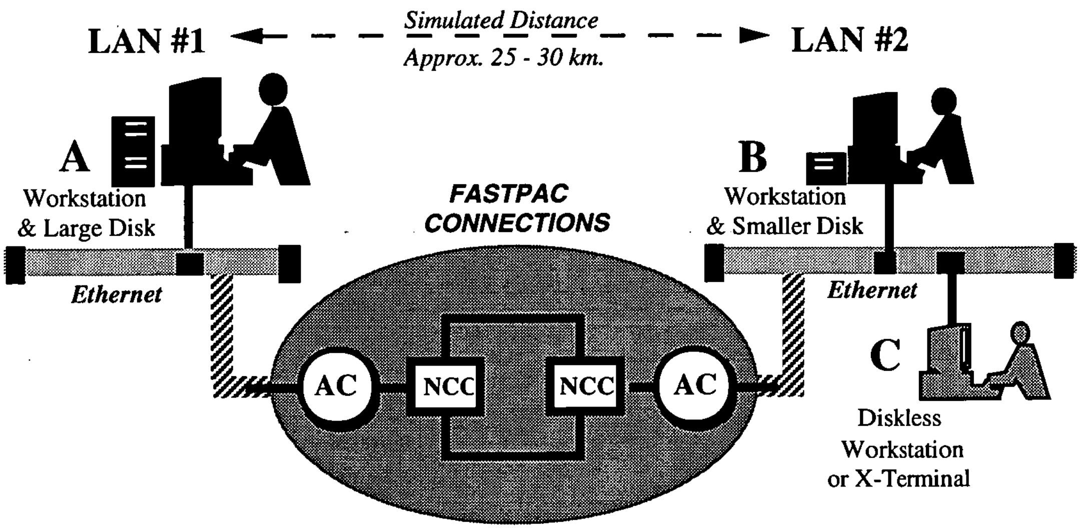


Figure 4.3: FASTPAC *FP[B]* Configuration

1992 Experiments

In the 1991 set of tests, Workstation *A* was configured as the Unix and NFS server for all three workstations with a view to testing the FASTPAC link under maximum loading conditions. However, in many organisations, network managers would more likely have Unix servers running on each LAN in order to minimise the amount of inter-network overhead traffic related to operating system commands and NFS server requests. In light of this, the network configurations in the 1992 experiments differed in three important respects:

(1) Workstations on both sides of the FASTPAC cloud operated full versions of the operating system, thereby reducing the amount of "housekeeping traffic" (i.e., low-level messages and acknowledgements to client workstations and terminals) travelling across the WAN gateway;

(2) Only the "longer" of the two FASTPAC configurations was employed; and...

(3) The experiments employed an *operating* 10 Mbit/sec FASTPAC service connecting Telecom Australia's offices in central Melbourne with the Computing Centre at Monash University -- a distance of approximately 20 kilometres. Key components of this configuration are illustrated in Figure 4.4.

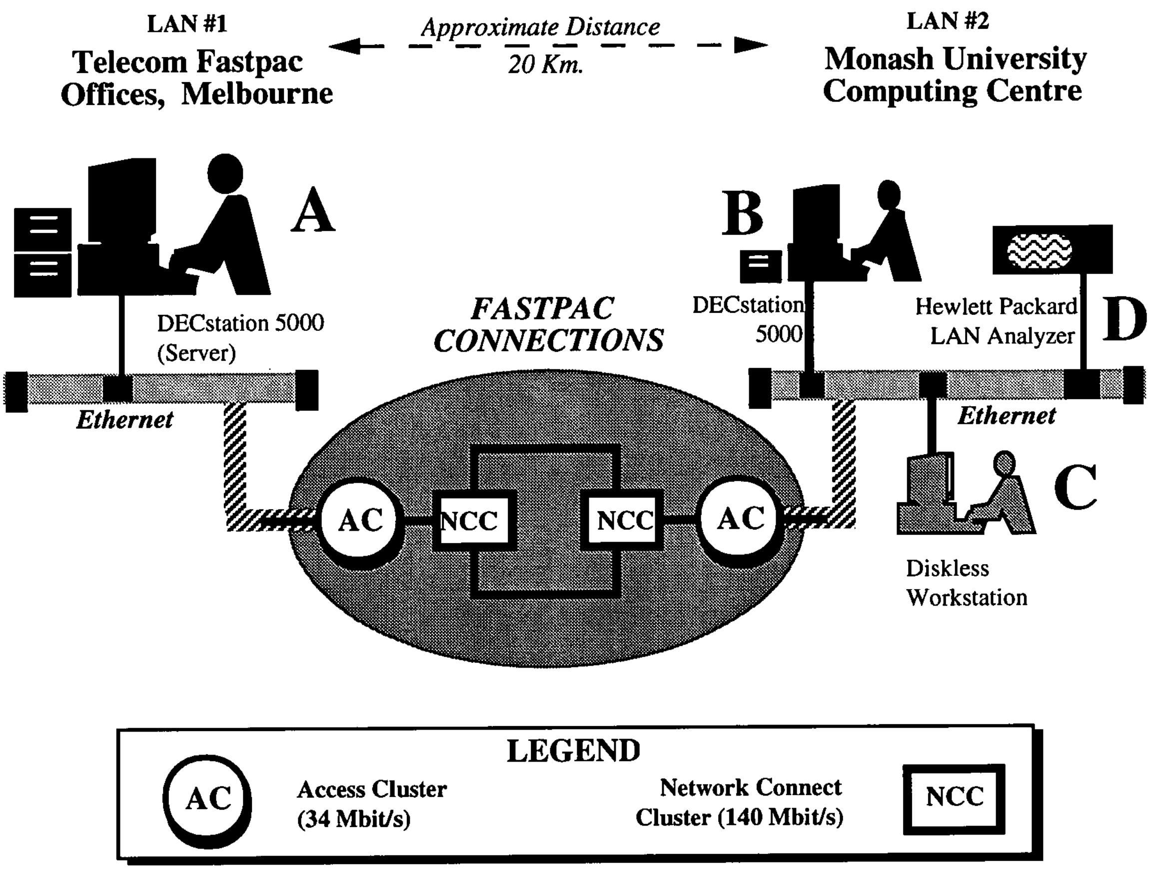


Figure 4.4

FASTPAC Network Configuration During Experiments

4.3 OPERATIONS AND DATA SETS USED IN THE TESTING

The specific operations employed in the testing are summarised in Table 4.3 and described briefly in the following two sections.

4.3.1 Unix File Transfer and Copy Operations

File transfer and copy operations are important and frequently-invoked activities within both local and wide-area networks. In addition to their employment by many end-users, bulk file transfer and copying activities are necessary for routine backup of data archives across a network — an important system management function. Unfortunately, due to the large file sizes involved in many GIS applications, users of traditional wide area networks have been reluctant to either transfer or back up these files between remote sites.

The UNIX "ftp" Operation

The *ftp* (or "file transfer protocol") command set is an application-level program within UNIX used to transfer a copy of a selected file or files directly from one workstation to another. An *ftp* operation is *explicit* in nature: i.e., the process involves identifying and logging in to the remote machine, switching to the desired directory; and then either: (a) retrieving a specified file (or files) stored on the remote workstation to the local machine (using a GET command); or (b) transferring a specified local file or files to the remote workstation (using a PUT command).

The UNIX "cp" Operation

On stand-alone workstations, the UNIX *cp* command is a means of copying a specified file into either the same or a different directory using the command syntax: ***cp*** */pathname1/file1 /pathname2/file2*

Table 4.3: Summary of Operations Employed in FASTPAC/GIS Testing

Test	Command	Specific Action	Nature of Data Traffic
GIS File Processing & Copying	*CLEAN (ARC/INFO)*	Build polygon topology for map-sheet coverage and clean any arc overshoots or undershoots. *(I/O and CPU-Intensive)*	• Entire input file loaded from designated NFS-mounted directory into memory of user's workstation; • Processing carried out at user's workstation; • Output file written incrementally to a second designated NFS-mounted directory.
	UNION (ARC/INFO)	Overlay two mapsheet coverages to form a combined set of graphics and attribute files of higher complexity and larger size. *(I/O and CPU-Intensive)*	Same as above
	COPY (ARC/INFO)	Copy coverage (graphics files, attribute files and database pointers) from one workspace to another.	• Entire input file loaded from NFS-mounted directory into local memory; • Copy of input file renamed and written to a second NFS-mounted directory.
GIS Display	*DRAW (ARCEDIT)*	Drawing arcs, nodes and label points of elected coverage.	• Data from 3 separate input files loaded from NFS-mounted directory into memory and displayed.
	POLYGONS (ARCPLOT)	Plotting boundaries of polygons possessing specific attributes	• Coordinates for a selected portion (60%) of original polygons loaded from selected NFS-mounted directory into local memory and displayed on screen.
	POLYGONSHADES (ARCPLOT)	Colouring polygons based on respective attribute values.	• Colour-code information for respective polygons sent to user workstation.
	IMAGE (ARCPLOT)	Displaying raster image	• Entire input file loaded from selected NFS-mounted directory into local memory; • Bit-map version of file displayed on screen
File Transfer	UNIX *cp* command	Copy from local disk to NFS-mounted directory on remote disk, and vice versa. (Employs UDP protocol)	
	UNIX *ftp* command	Transfer to and from local disk following *ftp* login to remote CPU. (Employs TCP/IP protocol)	

...where *pathname1* and *file1* are (respectively) the directory path and name of the original file and *pathname2* and *file2* are (respectively) the intended directory path and name of the copy.

Provided the respective directories are mounted and recognised by the NFS operating across a network, a file can be copied from a directory on one machine to a different directory on another workstation using virtually the same command syntax as that shown above. The fact that the respective directories reside on different machines is transparent to the end-user. Unlike the explicit activities necessary with the *ftp* command, then, the *cp* command (with the help of NFS) offers a more *implicit* means of file transfer between UNIX machines.

In these experiments, image and graphics files would be transferred or copied between the server *A* and the client *B* (in both directions) across different LAN and FASTPAC configurations using the Unix *ftp* and *cp* commands. The respective execution times would be measured for both these operations over all three network configurations ([i.e., LAN, short-distance FASTPAC (or "FP[A]"), and longer-distance FASTPAC (or "FP[B]")].

4.3.2 GIS Processing and Display Operations

Section 3.3 summarised the rationale behind the GIS commands chosen for use in the testing. The *ARC/INFO CLEAN*, *UNION* and *COPY* commands are routinely used during data loading and analysis processes, and have the potential to place a heavy load on the server and, in certain cases, on the network. The ARCEDIT *DRAW* and Arcplot *POLYGONS* and *POLYGONSHADES* commands represent operations in editing and database query sessions respectively where fast transfer and display of graphics is required. Finally, the Arcplot *IMAGE* command was also selected to determine the comparative times involved in displaying large image files across local and wide area networks.

4.3.3 Data Sets Employed

Table 4.4 summarises the characteristics of the data files used in both rounds of testing.

A combination of "real-world" and synthetically-generated data sets was used in both rounds of testing. Rather than using actual polygon data from one of the participating organisations, synthetic *ARC/INFO* coverages consisting of a regular grid of specified dimensions, spacing and orientation were generated for use in most of the GIS operations. For the ARCPLOT *POLYGONS* and *POLYGONSHADES* operations, which displayed selected polygons based on a specified selection criteria, single-digit integers between 1 and 9 were randomly assigned as "attributes" to the polygons in a manner similar to that described in [Marble et al., 1989].

The decision to employ synthetically-generated coverages for many of the operations was prompted by the desire to maintain experimental control over the size and density of the GIS data used in the testing. There is an important tradeoff in making this decision, since these regular gridded coverages will not be as complex as those found in real-world applications. The degree of data complexity (i.e, the number and shape of arc segments in each polygon) has a measurable effect on GIS performance: other research has demonstrated that *ArcInfo* performs certain operations faster using synthetic data sets than with more complex, real-world data coverages of the same size [Hawke, 1991b].

It was originally assumed that relative file size was a key determinant to data transfer performance and that the relative complexity of the arcs contained within these files would not affect the *relative* differences in response time between the same operations executed on different network configurations. With this in mind, only synthetic data sets were employed in the initial experiments. However, subsequent results indicate that data complexity may indeed have some bearing on selected NFS operations. This will be discussed further in Chapter 5.

In addition to the synthetically-generated files, a 9.8 Mbyte compressed LANDSAT image ("Testfile") was copied or transferred between workstations using the UNIX *cp* and *ftp* operations. As well, a 6 Mbyte scanned image of a black-and-white photograph ("Port") was used as the candidate data coverage for the ARCPLOT *IMAGE* operation.

Table 4.4: Data Sets Used in the FASTPAC/GIS Performance Testing

NAME	TYPE	SIZE	CONTENTS	COMMENTS
T50-G1	Vector	Total: ~722 Kbytes†	• 2500 polygons • 5096 arcs • corresponding text files	*Input data coverage* synthetically created by generating a 50x50 grid. Polygon attribute file contains one new field containing an integer value between 1 and 5. (Used with both Processing and Editing operations.)
T50-D1	Vector	Total: ~768 Kbytes†	• 2612 polygons • 5516 arcs • corresponding text files	*Input data coverage* synthetically created by generating a 50x50 grid. Similar to T50 -G1, but oriented at a 30° angle to the horizontal. (Used with Processing operations only.)
T50-U1	Vector	Total: ~3559 Kbytes†	• 11,126 polygons • 22,448 arcs • corresponding text files	*Generated data coverage* created by overlaying T50-D1 onto T50-G1. (Used with testing of ArcInfo *COPY* operations only.)
TESTFILE	Image	9.8 Mbytes±		Compressed satellite image. (28 Mbytes uncompressed). (Used with UNIX file transfer operations only.)
PORT	Image	6 Mbytes		Scanned image of black & white aerial photograph.

† *ARC/INFO "Coverages" are each actually directories consisting of 12 different graphics and text files.*

4.4 LOGICAL NETWORK USAGE CONFIGURATIONS

Previous network testing [Alexander and Fox, 1992] suggested that the predominant types of GIS file transfer and file sharing activities within an organisation may vary depending on prevailing attitudes towards two criteria:

(1) centralisation vs. de-centralisation of processors and applications; and

(2) centralisation vs. de-centralisation of data storage

Table 4.5 presents a framework based on the above criteria and suggests typical GIS- or network-related applications which may fall into each category.

Table 4.5: Potential GIS-Related Usage of Broadband Communications Under Different Information Management Approaches

Approach	Degree of Centralisation in...		Possible GIS-Related Usage of High-Speed Communications Network
	Processing	Storage	
1	Decentralised	Decentralised	• "Browsing" of remote files prior to retrieval; • High-speed file transfer between sites.
2	Centralised	Centralised *(All data stored on central file server)*	Any activities possible through network file sharing, including... • Copying of selected files to local workspace; • Processing, edit, analysis and display of centrally-stored files.
3	Centralised	Decentralised *(Local storage possible; only corporate data stored on central server)*	• Query and display of centrally-stored "corporate data"; • Transfer of selected files to local workspace for processing.
4	Centralised	Centralised	• Display and edit (i.e., re-drawing) processes would comprise the major load on the communication network.

Employing a similar two-factor framework, Gunton [1989] suggested that the relative importance applied to (1) *Rapid access to corporate data* and (2) *Local autonomy in processing* can be used to characterise different approaches to information management in an organisation. Using this criteria, he was able to suggest alternative information processing strategies most appropriate for an organisation with a given set of operational requirements.

Using this same classification scheme, Figure 4.5 illustrates how users in different organisations may assess whether network performance is "satisfactory" or not. For example, in an environment where both data storage and processing are centralised -- i.e., most users would access a central computer or server via remote login with interactive terminals -- users will gauge "network performance" holistically in terms of data display speeds. In other words, if response times and/or data display speeds are long when working at a remote terminal, then "the network" must be slow.

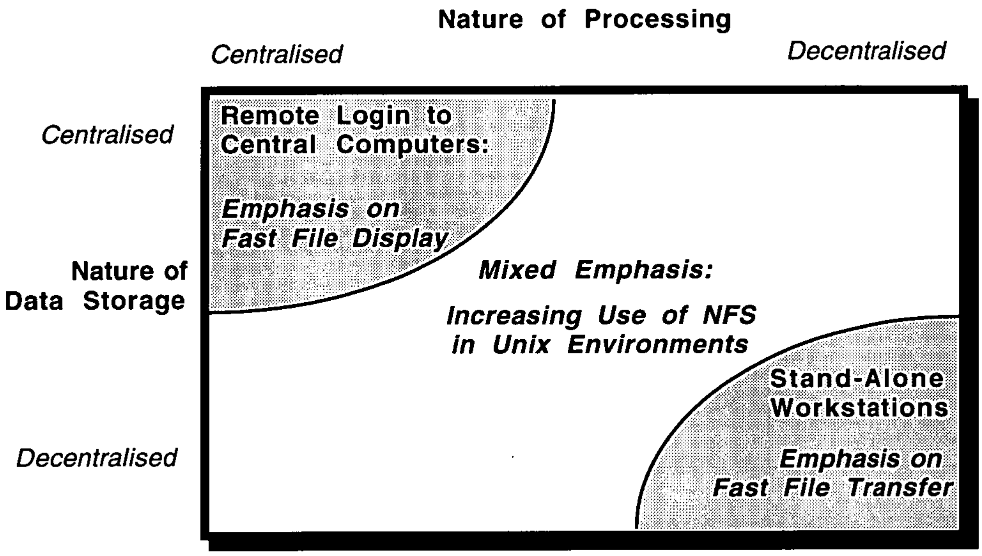


Figure 4.5
Networking Performance Emphasis in Varying Information Management Environments

In a <u>de</u>centralised environment, everyone relies largely on their own workstation or PC for standalone processing and interactive work. In certain cases, such users may often share large GIS or DBMS software resources within a LAN environment. However, most applications will still dictate the transfer or "downloading" of data to the local machine for processing and display.

"Network" performance in this context, then, may be gauged by users largely on the basis of file transfer speeds.

It is the vast middle ground where the fastest growth is now taking place in the LAN community. An increasing number of LAN users want to organise their resources into a virtual file system managed by NFS (or "NFS-type") networked resource management systems. (See Section 2.4.3 for details.) Using NFS, users may then transparently read and write data to disks which may be stored on their own machine, on a server in the next room, or on a larger system in a different building. Because transparency is the key advantage to this configuration, users tend to assess "network" performance in terms of interactive response, data display and file transfer rates and will compare them (unfairly, but understandably) to those experienced in stand-alone mode.

Given these varying requirements, the experiments themselves were designed to examine performance in three different categories of network usage. These categories included: (a) bulk file transfer; (b) remote login to a central file server; and (c) remote NFS-mounting of disks located elsewhere on the wide-area network.

Clearly, overall response time is based on more than just network-related factors. Dowers et al. [1991], Katz [1991] and others have already suggested that the hardware characteristics and load of disk and processor units can have a greater impact on performance in a client-server environment than network traffic and limitations. Even so, the framework is useful for high-level identification of the general types of operations to be employed in subsequent testing.

The third category mentioned earlier deserves special attention. Given the use of FASTPAC as a public telecommunications service with many customers sharing a common bus, centralised versus distributed network models will be realised *logically* but not *physically*. FASTPAC is designed primarily for LAN-to-LAN communication. In the Client-Server architectures implicit in these LANs, it is the respective locations of *the user, the host processor, the application software,* and *the input and output data* that will determine the logical configuration of the network for many applications.

If it can be assumed that the user will be operating at a workstation during this testing, then the first two of the above components can be merged. This assumption made, a series of logical *user-application-data* configurations was established by varying the respective geographic locations of the user, the disk containing input data files and the disk designated to contain the output files. The practical implementation of these various configurations was accomplished through the use of files strategically located (or specified for location) on different local and remotely-mounted directories which could be accessed transparently by UNIX and *ARC/INFO* using Sun's "Network File System" (NFS).

4.5 DEVELOPING TEST PROCEDURES

4.5.1 Preparation of Test Scripts

All operations were carried out through batch processing scripts prepared using *ARC/INFO's* Arc Macro Language (AML). In each script, operations were sequenced such that a new input data file had to be loaded into the workstation's memory before each operation. Different scripts were prepared to reflect the various possible configurations of network usage. Tables 4.6 and 4.7 contain summaries of the usage configurations employed in (respectively) the 1991 and 1992 rounds of experiments.

System utilities were employed to record execution and cpu-times, while a Hewlett-Packard LAN Analyzer was used to gather statistics on data traffic characteristics and to generate specified background traffic loads.

ARC/INFO and Unix file transfer tests were conducted separately. To facilitate the use of system performance-timing utilities, the operations were carried out through batch processing scripts prepared using the Arc Macro Language (or AML) for *ARC/INFO* operations and shell-script commands for the Unix *cp* and *ftp* operations. Different scripts were prepared for eight different possible configurations of network usage, including two which involved concurrent usage of network resources. All the AML scripts used in the testing are included in Appendix B.

4.5.2 Estimating Sample Size

AML scripts were written such that these ordered combinations of operations would be repeated a designated number of times to ensure the resulting estimates of mean response time fell within a desired limit of error.

Table 4.6: Usage Configurations Employed in 1991 GIS Network Testing

Priority	User @	App. @	Data @	Comments
1	A	A	A	Stand-alone workstation.
2	B	A	B	Application software resides centrally; data held locally.
3	B	A	A	Remote terminal into central CPU, but with local swap space.
4	C	A	A	Simulates remote terminal into central CPU and disk.
5	C	A	B	Simulates remote terminal in file sharing application with distributed data.
6	A	A	B	Application is local but corporate data is managed centrally.
7	B and C	A	A	To determine effects on bandwidth utilisation and performance when more than one user is on the network sharing the same disk.
8	B and C	A	A and B	To determine effects on bandwidth utilisation and performance when two users are on the network using different disks at different times.
LEGEND A = Sun IPC B = Sun IPC C = Diskless Sun IPC				

Table 4.7: Usage Configurations Employed in 1992 GIS Network Testing

Priority	User @	App. @	Data @	Comments
1	A	A	A	Stand-alone workstation.
2	B	A	B	Data held locally; Application software stored centrally.
3	B	A	A	Remote login to central CPU; local swap space.
4	A	A	B	Application software local; corporate data managed centrally.
5	C	A	A	Application software and all data resides on remote server; users log in using X-Terminal.
6	C	A	B	Application software resides on remote server; Data resides on separate disk at user's site; users log in using X-Terminal.
LEGEND A = DECstation 5000 B = DECstation 5000 C = Labtam X-Terminal				

The formula defining the sample size n required to estimate mean $\bar{x}$ within an estimated error bound of B can be expressed as follows:

$$n = \frac{Ns^2}{(N-1)D + s^2}, \text{ where: } D = \frac{B^2}{K^2}$$ (from Mendenhall et al., 1970])

where:
n = Estimated size of sample required
N = Size of entire population
s^2 = Largest observed variance in a sample
B = Desired limit of error for estimated mean
K = Multiple of standard deviation error selected to achieve specific degree of confidence (e.g., K=2 for 2σ, or 95% confidence)

For the purposes of this research, it was decided that the mean times required to perform GIS observations should be estimated to within an error bound of ±1 second for display-type GIS operations and ±5 seconds for more processing-intensive GIS operations. The relative magnitude of these numbers can be debated -- the differences simply reflect the author's belief that users will be more concerned with tight estimates of times required to complete interactive display operations. By comparison, users may be more tolerant of fluctuations in execution times for more processing-intensive GIS operations, since these may be carried out in the background or in off-peak hours.

Series of initial samples were taken and processed to provide estimates of the likely standard deviation *(s)* in execution times for various sets of operations. In those particular tests, the maximum standard deviations observed were ± 1.5 seconds for display-type operations like *DRAW*, *IMAGE*, *POLYGONS* and *POLYGONSHADES*, and ±5 seconds for processing-intensive operations like CLEAN, COPY and UNION. Using the above formulae with an estimated total population *N* of 1000, it was calculated that approximately 9-11 observations would be required to provide mean estimates which fell within the desired error bounds B 95% of the time. Fifteen to twenty observations were actually employed in the testing.

4.5.3 Minimising Bias

Preliminary experiments indicated that repeating each operation the designated number of times before moving on (i.e., *a-a-a...b-b-b...c-c-c*) and then taking the mean value of these operations generally *underestimated* performance times encountered in real-life operating conditions. Specifically, while a response time of x seconds would be observed on the first iteration, subsequent iterations would yield times of approximately y seconds, where $y < x$ (and sometimes $y << x$).

An example of this phenomenon found in early experiments is summarised in Table 4.8. After the first iteration, the response time for operation "CLEAN BAA" (which was originally intended to draw input data across the network from a remote disk) reduced to that of CLEAN AAA (the same operation in stand-alone mode). Since the user's workstation had drawn the necessary information into local memory and retained it there, the operation in subsequent iterations essentially changed from a *client-server* to *stand-alone* application.

Upon examination, this was found to be due to both the input data and the relevant portion of the GIS software object code being recognized and kept in memory between successive operations. This *memory cacheing* phenomenon — usually regarded as a positive feature of the Unix operating system — was observed to occur whenever the same single operation was performed on the same coverage(s) n times in a row. After the first iteration, the workstation would retain the requisite application code and data in memory and therefore would not need to draw that data from disk or across the network for subsequent repetitions. The corresponding execution times measured for the second and subsequent repetitions would therefore be shorter (since they would have no data transfer time components) and would bias the overall results.

Table 4.8: Observed Response Times Encountered when CLEAN Operation Repeated Individually

Iteration	CLEAN BAA (secs)	CLEAN AAA (secs)
1	79	52
2	54	54
3	53	53
4	54	53
5	53	53

Note: "CLEAN (*X-Y-Z*)" indicates an operation where required input files were read from disk at Workstation *X* and processed at the User's Workstation *Y*. Output files were then written to disk on Workstation *Z*.

To circumvent this problem, each AML script was redesigned to specify successive completion of *a set* of different operations (i.e., *a-b-c-d-e...a-b-c-d-e...a-b-c-d-e*) rather than repeating each operation the designated number of times before moving on (i.e., *a-a-a...b-b-b...c-c-c*). By ensuring that: (1) no two successive operations in the script sequence were the same; and (2) each operation employed different data, any effects due to the possible influence of memory cacheing would be minimised.

4.5.4 Dealing with Multiple Samples

Where logistics dictated that twenty iterations of the same script could not be completed in a single block, a test for a given configuration would be split into blocks containing a smaller number of iterations. Before grouping the results of these separate runs into a single sample, the *t-ratio* statistic was used to compare the respective mean value and standard deviation of each operation to determine if any significant difference existed between the corresponding samples. This statistic was calculated using the following formula:

$$t = \frac{\tilde{X}_1 - \tilde{X}_2}{\sigma_{\tilde{x}-\tilde{x}}}, \quad \text{where} \quad \sigma_{\tilde{x}-\tilde{x}} = \sqrt{\sigma^2_{\tilde{x}_1} + \sigma^2_{\tilde{x}_2}} \quad \text{and} \quad \sigma_{\tilde{x}_n} = \frac{\sigma_n}{\sqrt{N_n - 1}}$$

where:

$\tilde{X}_1 - \tilde{X}_2$ = difference between the means of the two samples;

$\sigma_{\tilde{X}-\tilde{X}}$ = standard error of the mean difference;

$\sigma_{\tilde{x}_n}$ = standard error of the mean of sample n, where n = 1, 2;

σ_n = standard deviation of the sample n ; and

N_n = number of observations in sample n.

4.6 POTENTIAL STRENGTHS AND LIMITATIONS OF THE DESIGN

The information collected to date and the testing methodology proposed in this chapter have been designed to obtain a reliable idea of the comparative differences in response-time performance between GIS-related applications in stand-alone and client/server configurations. Since the approach proposed employs operations and configurations used routinely, the results should provide the end-user with a better understanding of the performance tradeoffs involved. Moreover, provided the performance-timing utilities are available and can be supported by manual timing procedures, the basic approach and metrics employed are independent of both the hardware and GIS software in use.

However, there are limitations to this approach which must be understood when interpreting the results. These limitations include:

Specific hardware and software employed: Although the choice of performance metrics and design itself should be independent of the hardware and software employed, the actual performance results should be considered representative only of the particular combinations of hardware, network infrastructure and GIS software described in this document. Results obtained on these tests may be different if completed at a later date on more up-to-date equipment or later versions of UNIX and the *Arc/Info* software. More generally, the absolute performance times for the same operation on different systems will vary depending on the basic algorithms and the corresponding approaches to input/output and memory management strategies employed by different vendors. Nevertheless, discussions concerning underlying behaviour and general network patterns identified in the research should have a somewhat longer half-life, and general conclusions drawn from this research may be applicable to GIS operations found in many packages.

Access to a single type of Broadband Services: Due to logistical constraints, only one type of broadband service — the FASTPAC 10 service offered by Telecom Australia — was employed in this project. Purported network bandwidth is only one indicator of performance which can often be misleading when taken in isolation — different LAN interconnect services may possess characteristics which may influence performance on specific operations.

Controlled environment: Except where indicated, the methodologies employed and results obtained in the performance comparisons were based on an environment of controlled system and network loading with no special tuning of operating system or network parameters. Examining GIS performance in situations where individual users have full control over such parameters — while perhaps useful for system optimization and capacity testing — was considered to be beyond the scope of this particular research effort.

Limited Amount of Performance Space Actually Evaluated: Finally, this research design deliberately concentrates on breadth rather than depth in its assessment of GIS performance. As pointed out by Wagner [1991], Hawke [1991b] and others, determining the extent of GIS performance space on even one operation can involve testing a number of different datasets of varying densities under a variety of loading conditions. Extrapolating this to *completely* examine several different operations in a client/server environment using different datasets of varying complexity — while ensuring the results are statistically reliable — can consume more time and resources than those available for this research. While the approach proposed should give the user at least a preliminary idea of performance, recommendations for further research will be discussed in the final chapter.

A methodology for reliably and objectively comparing GIS performance across both a local and wide area network has been developed and the details of the actual testing procedures employed have been presented in this Chapter. The equipment, data, usage configurations and basic design considerations were described, and the potential limitations of the proposed approach have also been outlined. In the next chapter, the results of the actual performance tests are presented and analysed, and the practical implications of these results with respect to actual operating conditions are discussed.

5.0 PERFORMANCE MEASUREMENT: RESULTS AND ANALYSIS

In this chapter, the results of the FASTPAC/GIS performance measurement experiments are presented and analysed. The original purpose in conducting these experiments was to determine whether a broadband communication service indeed possessed the performance necessary to enable organisations with geographically dispersed sites to satisfactorily utilise and manage their geographic information system software and databases from a single location. While the results from still further testing would be required to confidently predict performance under all conditions, the results presented here do provide considerable insight into the behaviour and limitations of UNIX file transfer and *ARC/INFO* GIS operations in a client/server environment over broadband networks.

The first three sections in this chapter describe the results of performance measurement experiments on both loaded and unloaded networks. The first section describes the results when transferring and copying large files across LAN and FASTPAC metropolitan area networks with no other traffic load present. The second section examines and compares the response-time performance of selected *ARC/INFO* operations across the same (otherwise empty) media. Finally, the third section examines the effects of varying network and server traffic loads on the response-time performance of a limited number of *ARC/INFO* operations.

The final section discusses the contribution of these experiments to supporting or refuting the original hypothesis. In addition to assessing the extendability of these results to other operations, the section also examines the importance of response-time performance to end-user satisfaction in a client-server environment.

5.1 Performance on UNIX File Transfer and Copy Operations

This section examines the response-time performance involved in: (a) transferring a large image file between workstations using the UNIX *ftp* command; and (2) copying of GIS graphics files between remote, Network File System- (or NFS-) mounted directories on different workstations using the UNIX *cp* command. In both cases, the comparative times involved in performing these operations across a local area network and between LANs across the two FASTPAC links described in Section 4.2.2.

5.1.1 GENERAL COMMENTS AND OBSERVATIONS

As discussed in Section 4.3.1, the *ftp* (or "file transfer protocol") command set is an application-level program within UNIX used to transfer a copy of a selected file or files directly from one workstation to another. By comparison, the UNIX *cp* command is a means of copying a specified file into either the same or a different directory on the same machine, or even to a different machine using NFS.

There is a fundamental difference in the manner in which the data is packaged for communication in each case. The *ftp* operation makes use of the Transmission Control Protocol (TCP), while NFS — and, by extension, the UNIX *cp* operation — is built on top of the User Datagram Protocol (UDP). TCP was designed to provide reliable, sequenced delivery of packets over (relatively) long-lived network connections. UDP, on the other hand, is more of a "no frills" protocol which sends large datagrams to a remote host, but makes no assurances regarding their delivery or the order in which they arrive [Stern, 1991]. UDP is better suited for "connectionless" communication environments like NFS (i.e., where no explicit log-in is made) in which no context is required to send packets to a remote machine.

Although in theory the UDP protocol's simplicity means faster overall performance than that offered by TCP, the overheads and security precautions added by the operating system in general and NFS in particular mean that file transfers can take much longer using the *cp* operation than with *ftp*. This difference is illustrated in Figure 5.1, which compares the respective speeds of transferring a 9.8 Mbyte image file using *ftp* and copying the same file between machines using *cp*. A full breakdown of the respective response-time figures involved in UNIX *ftp* and *cp* file transfers may be found in Appendix C.1.

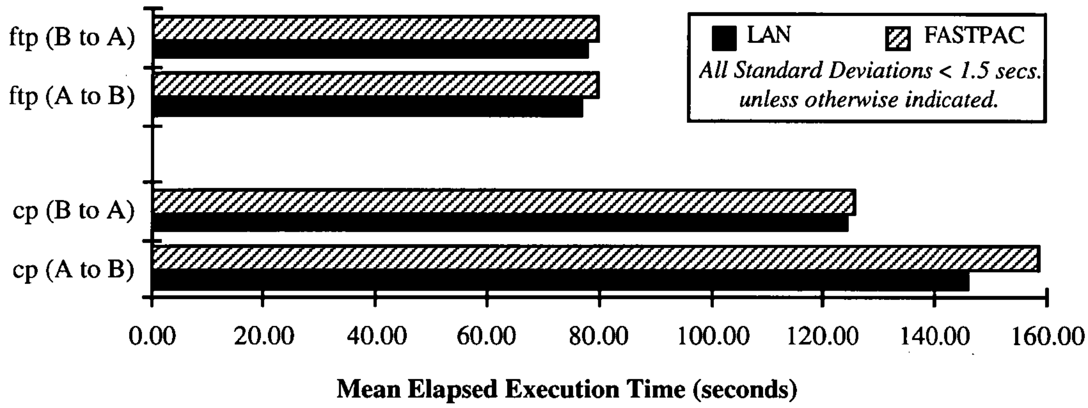


Figure 5.1
Execution-Time Comparison: UNIX *cp* and *ftp* Operations
(9793 Kbyte File)

Figure 5.1 illustrates the respective differences in response time across both a LAN and the longer of the two FASTPAC connections. The experiments in question were carried out at Telecom Australia's FASTPAC offices in Melbourne during October, 1991. LAN vs. FASTPAC performance differences notwithstanding, the bars in the above Figure clearly demonstrate the order of difference between the times required to complete the respective operations. Each of these operations will be discussed in more detail in following two sections.

5.1.2 UNIX *FTP* (FILE TRANSFER) OPERATIONS

In the first series of data transfer experiments, copies of a 9,793 Kbyte image file were placed on two separate workstations connected across: (a) the same LAN; (b) a short FASTPAC connection; and (c) a longer-distance FASTPAC connection. (See Section 4.2.2 for details.) The *ftp* GET and PUT operations were carried out on both workstations to observe any potential hardware- or command-related differences between receiving and transmitting operations on the two machines. The consistency demonstrated in each set of observations was very high, with no set possessing a standard deviation of greater than ±1 second.

In most cases, response-time differences between FASTPAC and LAN transfer times were minor. (See Figure 5.2.) In the two cases where data flowed from the Client workstation (B) to the Server (A), statistical t-tests confirmed that LAN vs. FASTPAC differences were within the noise levels of the respective sets of observations.

Slightly larger increases in response times across FASTPAC *were* observed in cases where data flowed in the other direction (i.e., from A to B). On reflection, these differences may be due to either: (a) the extent to which the files were fragmented on each disk; (b) performance constraints on A due to its double purpose as the UNIX server for the network; or (c) some other unidentified source of delay along that particular data path. However, since even these differences represented only a 3-4% increase over the corresponding LAN times, they were judged to be insignificant in the context of most end-user requirements and were therefore not examined any further.

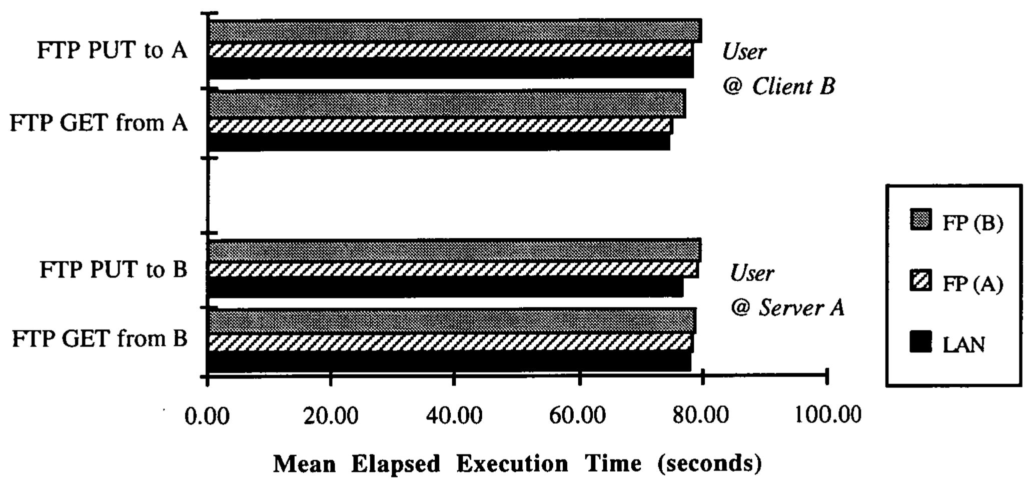


Figure 5.2

Execution Times for FTP Operations ***(9793 Kbyte File)***

5.1.3 UNIX *cp* (NFS COPY) OPERATIONS

Two files were used in this testing (3559 Kbytes and 9793 Kbytes respectively). These files were copied internally on each workstation to provide a baseline performance figure, and then copied — in both directions — between NFS-mounted directories on the Unix Server (A) and the Client workstation (B). Results from these tests are illustrated in Figures 5.3 and 5.4 respectively.

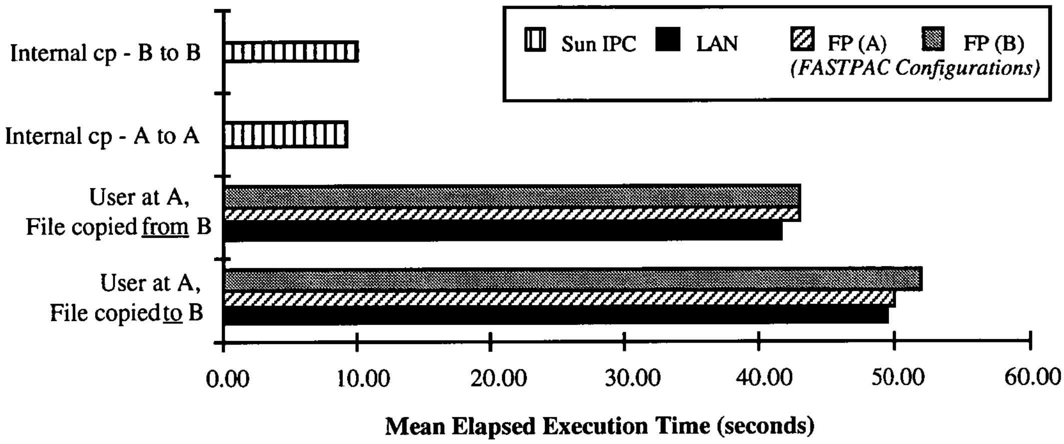


Figure 5.3
Execution Times for Unix *cp* Operations ***(3559 Kbyte File)***

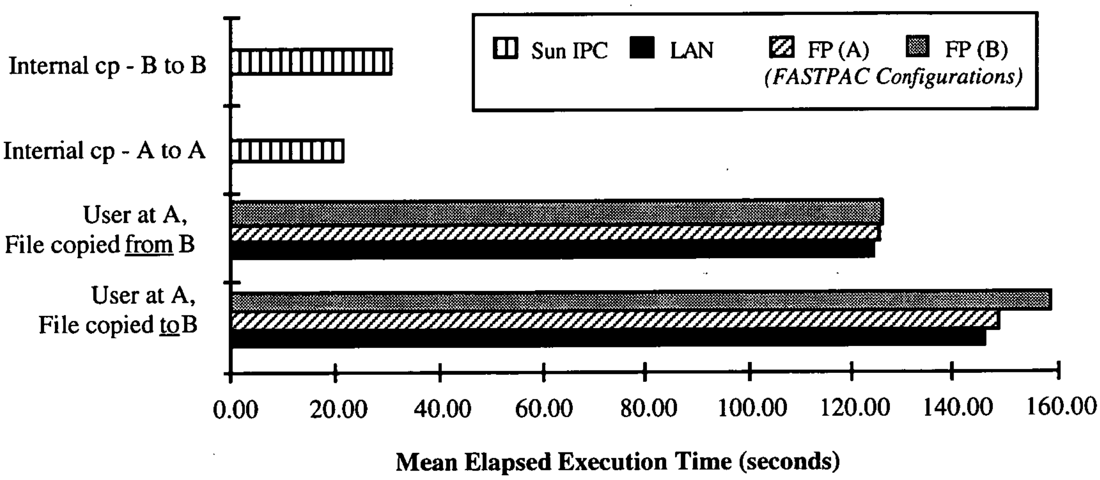


Figure 5.4
Execution Times for Unix *cp* Operations ***(9793 Kbyte File)***

As with the *ftp* operations, individual observations once again demonstrated close agreement, with no standard deviation on any set exceeding ±1.5 seconds.

As can be seen in both Figures, there is a marked difference in performance when one uses the *cp* command as a means of *file transfer* between machines rather than for simple file duplication on a stand-alone workstation. Copying a file between NFS-mounted directories generally took 4-7 times longer than copying the same file internally. More significant, copying data *from* the local machine (A) *to* the remote workstation (B) took much longer to complete than those where data was copied from the <u>remote</u> machine into the local workspace. Incremental delays introduced by FASTPAC were larger as well. This phenomenon will be discussed further in Section 5.2.3.

Finally, internal copies (i.e, in stand-alone mode) took significantly longer to perform on the Client than on the Server. This may perhaps be the result of either the actual data writing rate on one of the Server's disk drives (although the same models) being measurably faster than the Client workstation's internal disk drive (which is possible but unlikely) *or* additional overheads caused by the Unix operating system being supplied from a remote server.

5.1.4 IMPLICATIONS OF UNIX FILE TRANSFER AND COPYING RESULTS

While some degradation was observed in *copying* files under certain configurations, the results demonstrated that UNIX *file transfer* operations can effectively be performed at LAN speeds across FASTPAC. Users should be able to transfer files between their workstation and distant file servers as routinely as they now move data around the local area network at a single location. In practical terms, this can potentially allow routine and independent usage of large image and graphics files by remote end-users while still permitting the organisational benefits offered by centralized data storage and management.

5.2 Performance on GIS Display and Processing Operations

This section discusses and compares the respective times required to complete selected GIS display and processing operations in different "*User-Application-Data*" configurations for stand-alone, LAN, FP[A] and FP[B] options. A full breakdown of both the 1991 and 1992 experiment observations and results may be found in Appendices C.2 and C.3 respectively.

5.2.1 SELECTED GIS OPERATIONS -- 1991 EXPERIMENTS

Display Operations

The ARCEDIT *DRAW* command is used to display selected types of features and is commonly used in the edit and preparation of GIS map files. By comparison, the ARCPLOT *POLYGONS* command is invoked commonly by end-users to display (but not edit) selected graphics files. Similarly, *POLYGONSHADES* is a command which colours in specific polygons according to values of specific attributes stored in the GIS database.

In both the 1991 and 1992 experiments, the arcs, nodes, and labels for coverage *T50-D1* (2612 polygons, 768 Kbytes) were DRAWn. In related experiments, the POLYGONS and POLYGONSHADES operations plotted and shaded 80% of the 2612 polygons in the same coverage. In all these operations, the data was transferred (via NFS) from the disk of the workstation on which it was residing into the memory of the user's machine.

Figures 5.5 to 5.7 summarise the time differences observed when invoking the three GIS display operations selected for the 1991 experiments across different usage configurations. General observations include:

- *Consistent results*: In all cases where there was no background traffic load, the calculated standard deviations for each sample (20 observations/sample) were less than ±1.5 seconds.

- *Small but measurable delays introduced by the FASTPAC link.* Particularly with the shorter POLYGONS and POLYGONSHADES operations, the absolute differences between corresponding operations across LAN and FASTPAC were often negligible. In cases where the differences *were* found to be statistically significant, the times required to display data across the FASTPAC link were consistently longer than those observed across a single LAN.

- *Slightly longer response times on diskless workstations.* While statistical t-tests indicated that many of the differences fell within the respective "noise" levels of the samples, observed response times were consistently longer again in cases where data was displayed on the diskless workstation C. Further, the incremental increases in response times across the FASTPAC link were disproportionately larger for these operations than in those involving display on the disk<u>full</u> workstations.

In no instances did the incremental delays experienced under FASTPAC increase the response times over those conducted on a LAN by more than 20% (on diskless workstations), and in most cases the relative increases were less than 10%. The highest relative increases across FASTPAC were usually observed in configurations where the data was loaded and displayed on diskless workstations. While the file itself was small, additional data traffic may have been generated by swapping requisite data back and forth as required between the workstation's memory and its disk "scratch space". Since the allocated scratch space was located on the remote sever, these increases in response time appear to be the cumulative result of incremental delays introduced by the respective network routers and gateways at each end of the connection.

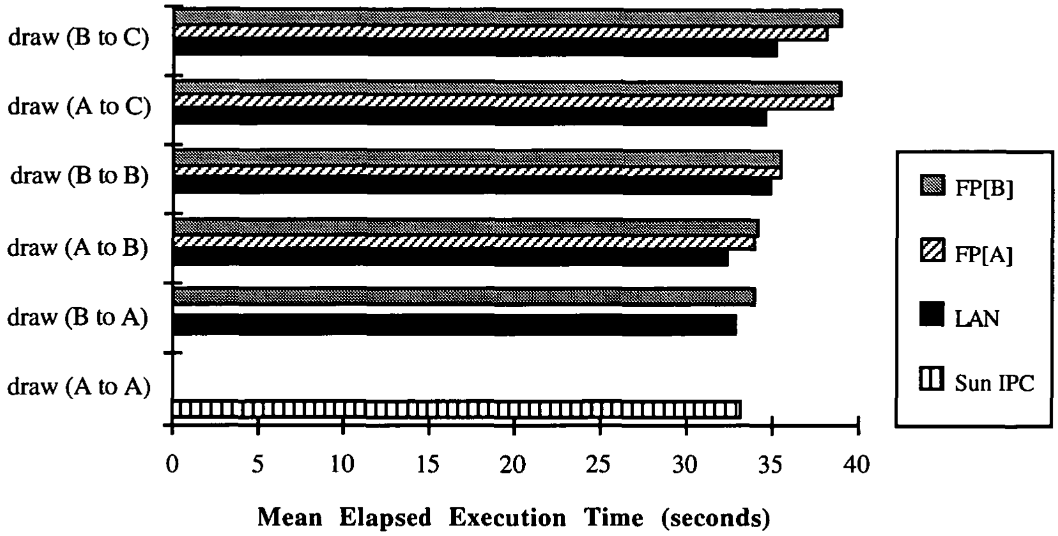


Figure 5.5

Execution-Time Comparison — ARCEDIT *Draw* Command

Note: "draw (*X* to *Y*)" indicates an operation where that the graphics files were sent across the network from Workstation *X* to be displayed on Workstation *Y*.

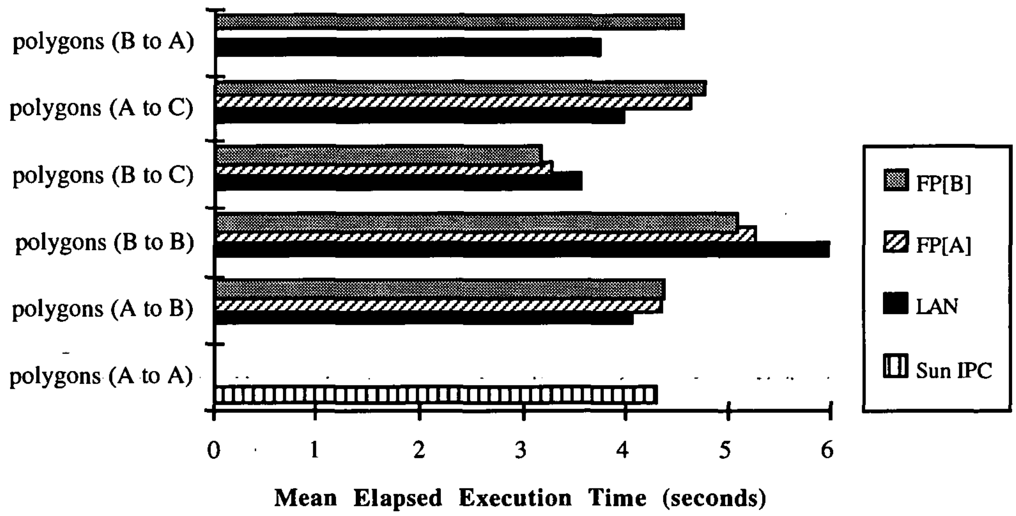


Figure 5.6

Execution-Time Comparison — ARCPLOT *Polygons* Command

Note: "polygons (*X* to *Y*)" indicates an operation where that the graphics files were sent across the network from Workstation *X* to be displayed on Workstation *Y*.

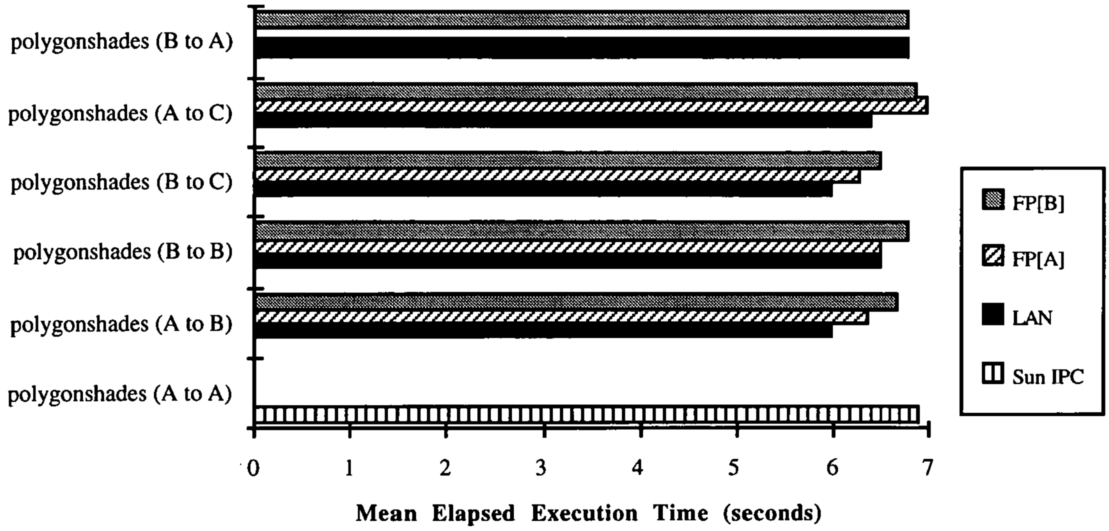


Figure 5.7

Execution-Time Comparison: ARCPLOT *POLYGONSHADES* Command

Note: "polygonshades (*X* to *Y*)" indicates an operation where that the graphics files were sent across the network from Workstation *X* to be displayed on Workstation *Y*.

One anomaly did appear with the POLYGONS operation when graphics files were pulled from Client B and displayed on either Clients B or C. (See Figure 5.6.) In these cases, it consistently took longer to display in the LAN configuration than across FASTPAC — even in cases where the files at B were being displayed on B. There is no immediate explanation for this anomaly. However, given that: (a) the absolute time required for the operation was small (~4-6 seconds); and (b) the differences fall within the acceptable error levels, this anomaly may not be significant.

Display of large graphics files and shaded polygon coverages on remote workstations can be a time-consuming task when working across lower-speed services. The results of these experiments suggest once again that — using FASTPAC — users on disk<u>full</u> workstations would finally be able to display remotely-stored files at virtually the same speed as those much nearer the file server. Users on diskless workstations would experience some delays, but these may be minimized by increasing memory or placing the UNIX server on the same LAN.

Processing- and I/O-Intensive Operations

Results from 1991 experiments examining response-time behaviour of the *COPY, CLEAN and UNION* commands are illustrated in Figures 5.8, 5.9 and 5.10 respectively. General observations include:

- *Consistent Observations:* While the respective samples of 20 observations each demonstrated a larger spread ($s \leq \pm 6$ seconds) than the GIS display operations discussed earlier, even the highest standard deviations amounted to only 5% of the estimated mean value.

- *Increased delays introduced when moving from stand-alone systems to a client/server environment:* Although there were exceptions, the same GIS operations usually took significantly longer to complete in a client-server environment. Depending on the operation involved and the respective locations of the input and output files, it would take anywhere from 10% to 60% longer to complete the operation across the LAN than on a stand-alone workstation.

  The only observed exceptions appeared when COPYing files from one disk to another, in which case it appeared as though the software had been optimised to begin writing data to the new file while still reading from the original, thereby allowing both disks to be utilised simultaneously. In these cases, the operation was actually completed faster across the LAN than on the stand-alone workstation.

- *Measurable delays introduced across the FASTPAC cloud:* Depending on the configuration, response times across the FASTPAC cloud were up to 25%

higher again than those observed across the LAN connection. Even in cases where the input and output data files resided on the same side of the FASTPAC cloud, the execution times increased under the FASTPAC options. However, while there were exceptions, the incremental increase in response times across FASTPAC were generally less than those encountered when moving from a stand-alone workstation onto a LAN.

- *Operations on diskless workstations take longer to complete.* Corresponding processing- and I/O-intensive operations carried out across a LAN generally took from 5% to 10% longer on diskless workstations than on diskfull workstations. As well, the incremental increases in response time across FASTPAC (as opposed to those observed on the LAN) were disproportionately larger on operations involving diskless workstations. Although the respective increases varied depending on the configuration, the trends were consistent: operations on diskfull workstations A and B increased by an average of only 6%, while those involving the diskless workstation C increased by an average of 11%.

- *Performance varied significantly depending on the respective network locations of the input and output data file in relation to the user's position.* Operations which read data from a remote NFS-mounted disk took marginally longer than those which read data from the local disk. However, operations which had to *write* data to a remote disk across the network took longer still. Finally, operations which had to both read *and* write data to remote disks took longer again.

- *NFS write-related differences were magnified across the FASTPAC cloud.* Operations which wrote data to a remote disk on the other side of the FASTPAC cloud took disproportionately longer to complete than those which either:

(a)wrote data to other disks on the same LAN or (b) read data from remote disks on either side of the FASTPAC cloud.

The implications of these observations will be discussed further in Section 5.2.3.

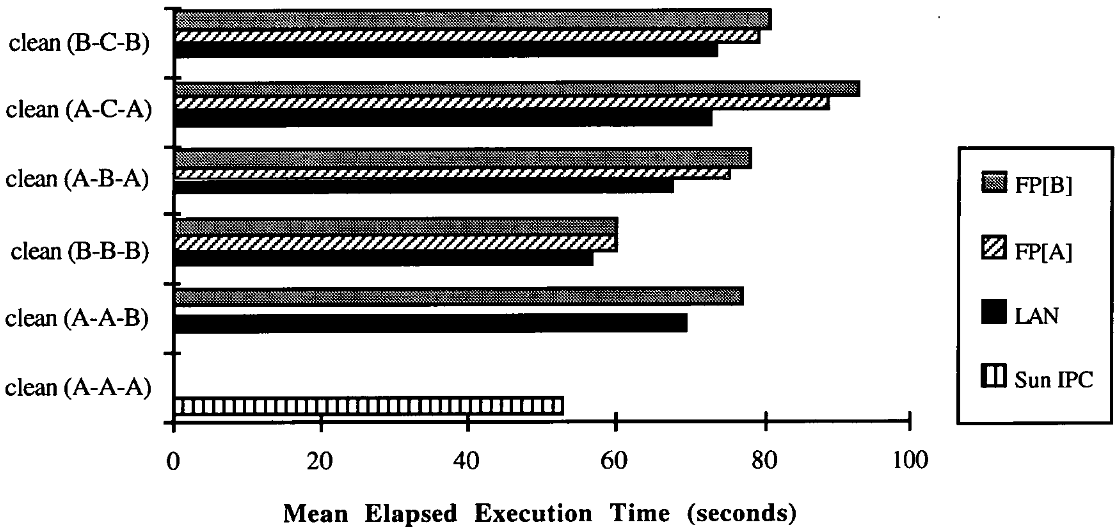


Figure 5.8
Execution Time Comparison — *ARC/INFO CLEAN* Command

Note: "CLEAN *(X-Y-Z)*" indicates an operation where required input files were read from disk at Workstation *X* and processed at the User's Workstation *Y*. Output files were then written to disk on Workstation *Z*.

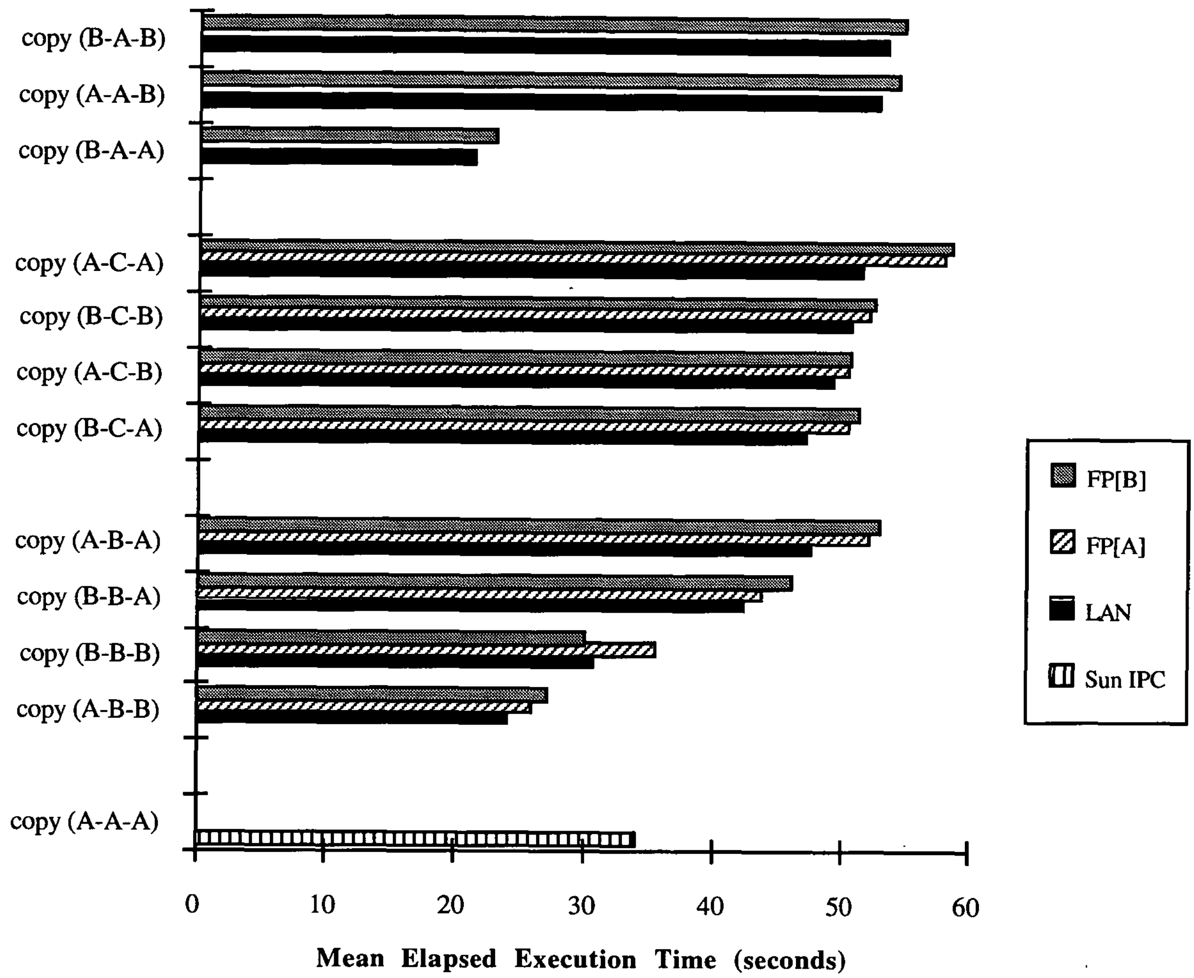


Figure 5.9

Execution Time Comparison — *ARC/INFO COPY* Command

Note: "COPY (*X-Y-Z*)" indicates an operation where required input files were read from disk at Workstation *X* and processed at the User's Workstation *Y*. Output files were then written to disk on Workstation *Z*.

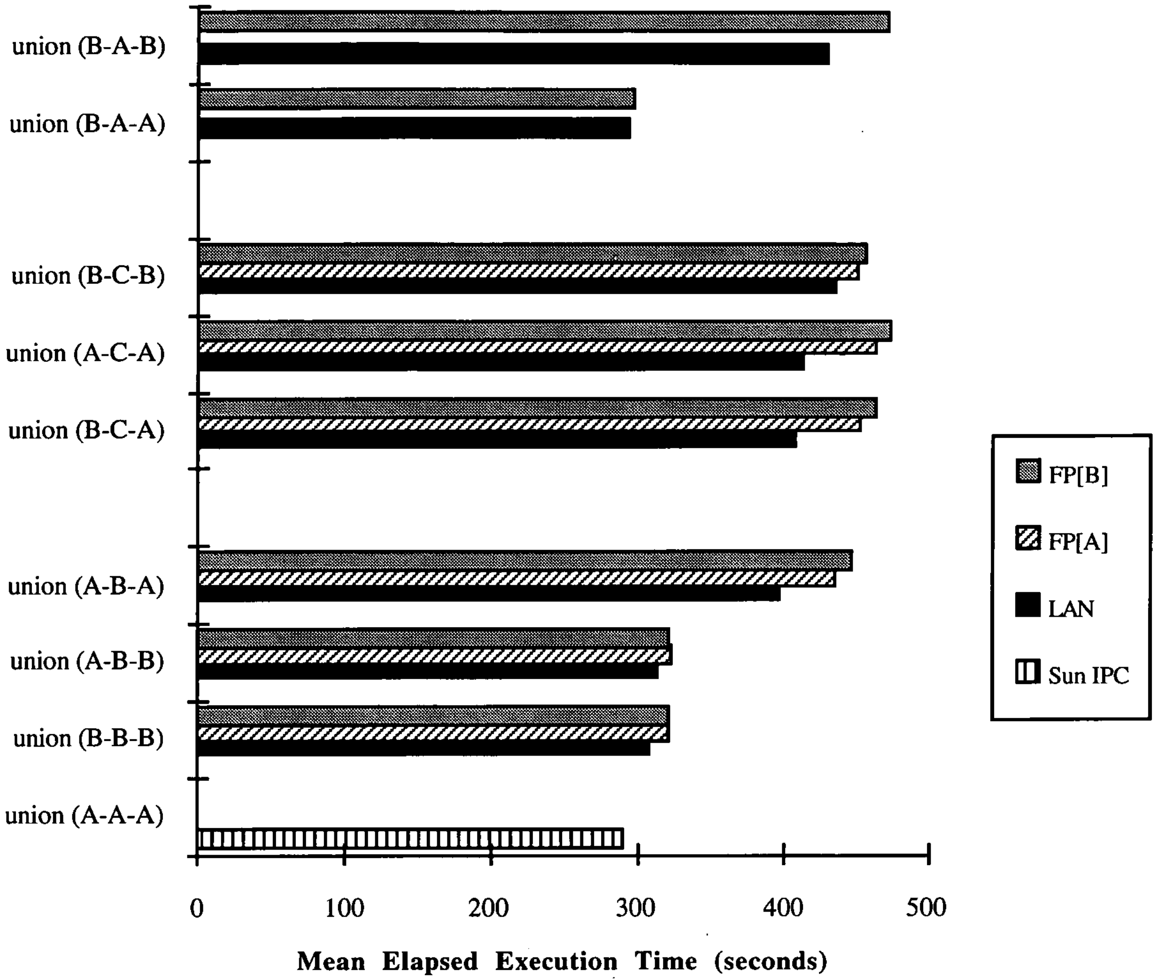


Figure 5.10

Execution Time Comparison — *ARC/INFO UNION* Command

Note: "UNION *(X-Y-Z)*" indicates an operation where required input files were read from disk at Workstation *X* and processed at the User's Workstation *Y*. Output files were then written to disk on Workstation *Z*.

5.2.2 SELECTED GIS OPERATIONS -- 1992 EXPERIMENTS

In the 1992 experiments, only the longer of the two FASTPAC connections was employed and identical diskfull DECstation 5000 workstations (labelled A and *B*) each possessed their own copies of the UNIX operating system. Rather than a diskless workstation, an X-Terminal (labelled *X*) was employed in this series of experiments. A full breakdown of the 1992 results may be found in Appendix C.3.

Display Operations

Figures 5.11 and 5.12 indicate the respective times required to complete *ArcInfo DRAW* and *IMAGE* operations in different usage configurations. There was very little spread among the 20 observations in each set once again, with the largest standard deviation being under ±0.85 seconds.

On average, *DRAW*ing the 768 Kbyte coverage *T50-D1* took only 2% to 3% longer to complete across FASTPAC than it did across a single LAN. As in the 1991 testing, the *DRAW* command seemed to perform at similar rates whether in stand-alone mode or across networks, and was largely unaffected by changes in network loading.

The scanned image was displayed using two different approaches. (See Figure 5.12.) For most tests, the command simply specified the display of an image file stored on a remote disk and NFS handled the arrangements by retrieving the entire file (~6 Mbytes) to the local system and displaying it on the screen. In at least one test, however, the IMAGE command was invoked from the X-Terminal logged into alternative remote host workstations. Under this configuration, *only the bit-mapped image required to fill the screen* (i.e., 1024 x 1024 pixels by 1 byte/pixel = ~1.05 Mbytes) was transferred across the network. As a result, the respective times required to display the image on the X-terminal in question were significantly less than those observed on the workstations.

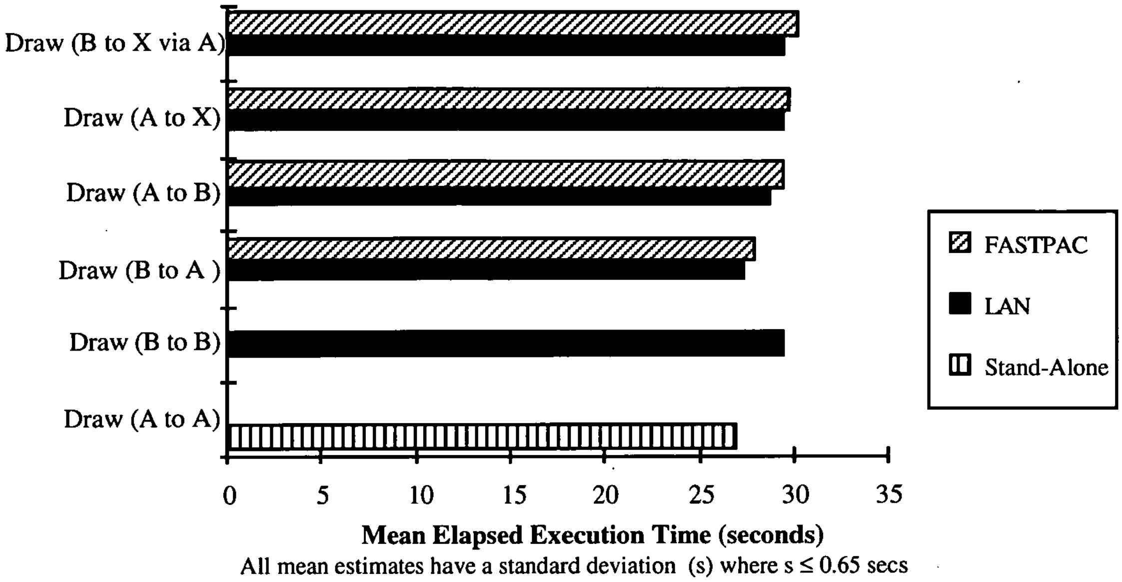


Figure 5.11

Execution-Time Comparison — ARCEDIT *Draw* Command

Note: "Draw (*Y* to *Z*)" indicates an operation where that the graphics files were sent across the network from Workstation *Y* to be displayed on Workstation *Z*. "X" indicates X-Terminal.

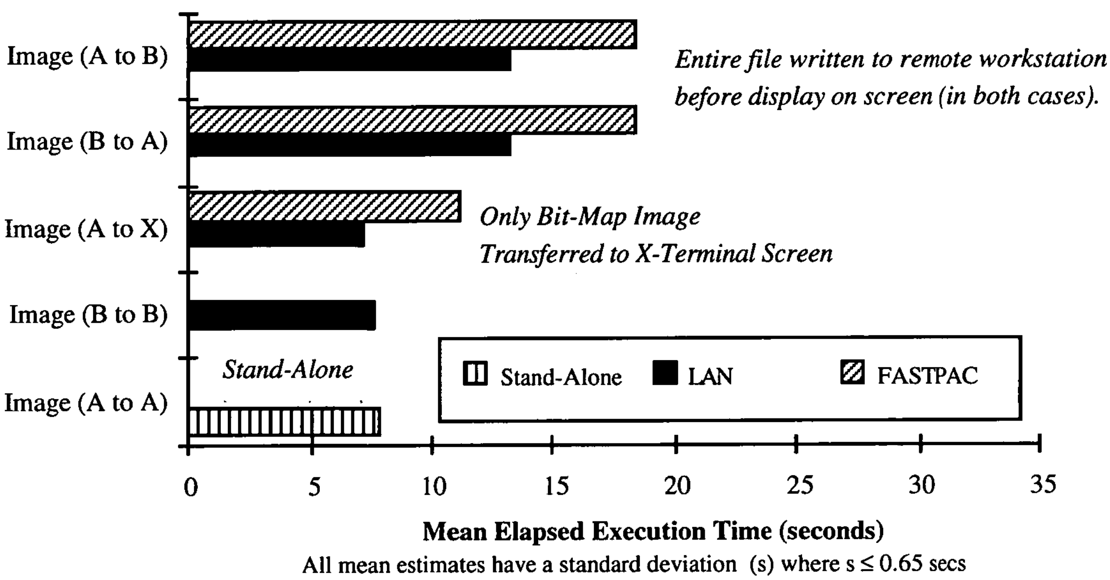


Figure 5.12

Execution-Time Comparison — ARCPLOT *Image* Command

Note: Except where indicated, "Image (*Y* to *Z*)" indicates an operation where the image file was sent from Workstation *Y* to be displayed on Workstation *Z*. (Note: "X" indicates X-Terminal.)

The image file itself took much less time than the graphics file to display, probably due to the lower hardware/software overheads involved in displaying bit-mapped images versus vector graphics on a high-performance raster workstation. However, the incremental delays introduced across FASTPAC were up to 40% higher than those encountered in the *DRAW* operation. This may be due in part to NFS-related considerations (which will be discussed further in Section 5.2.3), but also due to the sheer volume of data which must be transferred across the FASTPAC cloud.

Processing or I/O-Intensive Operations

Figures 5.13 and 5.14 (respectively) indicate the times required to complete *ArcInfo CLEAN* and *COPY* operations — in different usage configurations— on a stand-alone workstation, across a single LAN and between LANs across the 20 km. FASTPAC link. (The UNION operation was not tested in the 1992 experiments.) The observations demonstrated a high degree of consistency, with most mean estimates possessing a standard deviation of under ±1 second, and the maximum standard deviation being ±2.1 seconds.

Under optimal conditions, these particular operations took anywhere from 2% to 23% longer to complete across FASTPAC than they did across a single LAN, depending on the usage configuration involved. As before, client-server usage configurations which involved writing data to a remote disk took much longer to complete than their counterparts. As well, the incremental delays introduced across FASTPAC were disproportionately higher in these configurations.

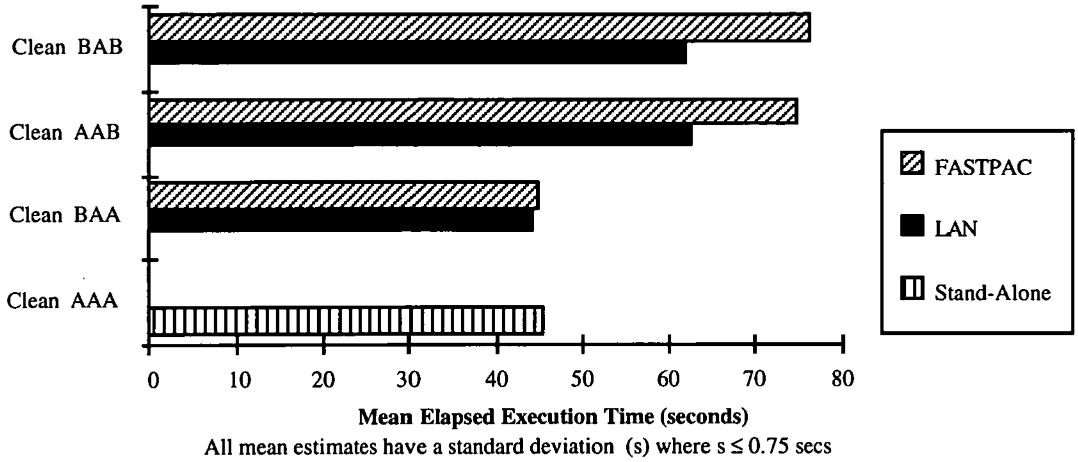


Figure 5.13

Execution Time Comparison — *ARC/INFO CLEAN* Command

Note: "CLEAN (*X-Y-Z*)" indicates an operation where required input files were read from disk at Workstation *X* and processed at the User's Workstation *Y*. Output files were then written to disk on Workstation *Z*.

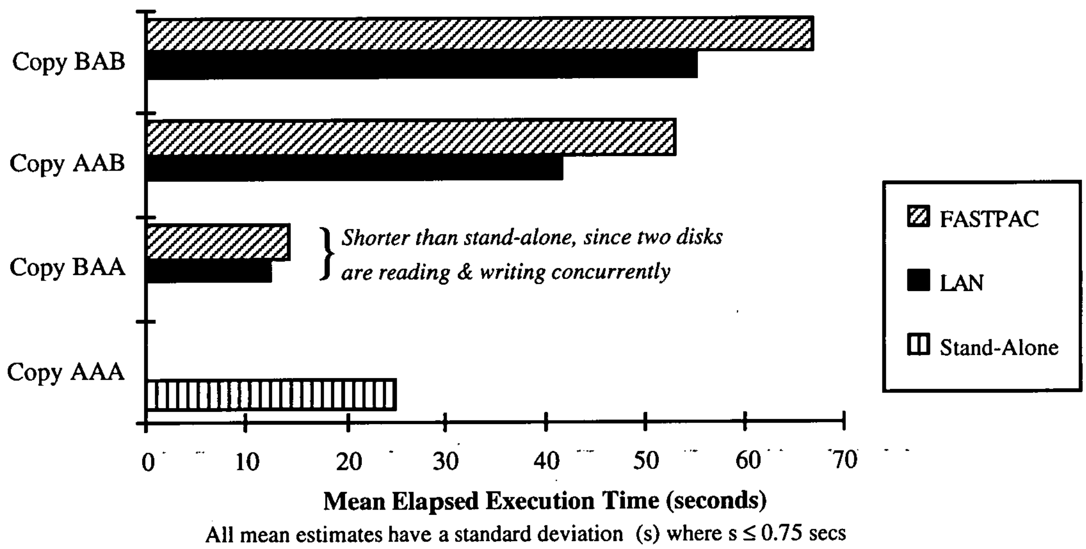


Figure 5.14

Execution Time Comparison — *ARC/INFO COPY* Command

Note: "COPY (*X-Y-Z*)" indicates an operation where required input files were read from disk at Workstation *X* and copied to the User's Workstation *Y*. Output files were then written to disk on Workstation *Z*.

Effects of Varying the Physical Location of the Software

In cases where the same operations were performed on comparable usage configurations (i.e., AAA vs. BBB; BAA vs. ABB; ABA vs. BAB; and AAB vs. BBA), the *t-ratio* statistic was used to compare the respective mean values and determine if any significant difference existed between the corresponding samples. This statistic was calculated using the following formula:

$$t = \frac{\tilde{X}_1 - \tilde{X}_2}{\sigma_{\tilde{x}-\tilde{x}}}\text{, where } \quad \sigma_{\tilde{x}-\tilde{x}} = \sqrt{\sigma^2_{\tilde{x}_1} + \sigma^2_{\tilde{x}_2}} \quad \text{and} \quad \sigma_{\tilde{x}_n} = \frac{\sigma_n}{\sqrt{N_n - 1}}$$

where:

$\tilde{X}_1 - \tilde{X}_2$ = difference between the means of the two samples;

$\sigma_{\tilde{X}-\tilde{X}}$ = standard error of the mean difference;

$\sigma_{\tilde{x}_n}$ = standard error of the mean of Sample *n*, where n = 1, 2;

σ_n = standard deviation of Sample *n* ; and

N_n = number of observations in Sample *n.*

The respective execution times for comparable samples in the CLEAN operation are illustrated in Figure 5.15, while a detailed summary of the *t-ratio* calculations for all operations is included in Appendix D.

A comparison of corresponding usage configurations indicates that — at least with the levels of memory possessed by workstations employed here — GIS performance generally remained similar *regardless of whether the application was stored locally or resided on the remote server across the FASTPAC cloud.* Differences in some corresponding mean values *were* interpreted to be statistically significant at the ".01" confidence level. Practically speaking, however, none of the corresponding times tested differed by more than 20% of the overall time involved and most differed by less than 8%.

This implies that, under low or sporadic loading conditions, remote GIS users could enjoy comparable levels of GIS performance as local users without having to obtain their own software. This has significant implications for large organisations who may operate many regional offices but still wish to manage their software centrally. Whether this is a practical alternative would depend on the specific customer's circumstances and the licensing restrictions of the application software in question.

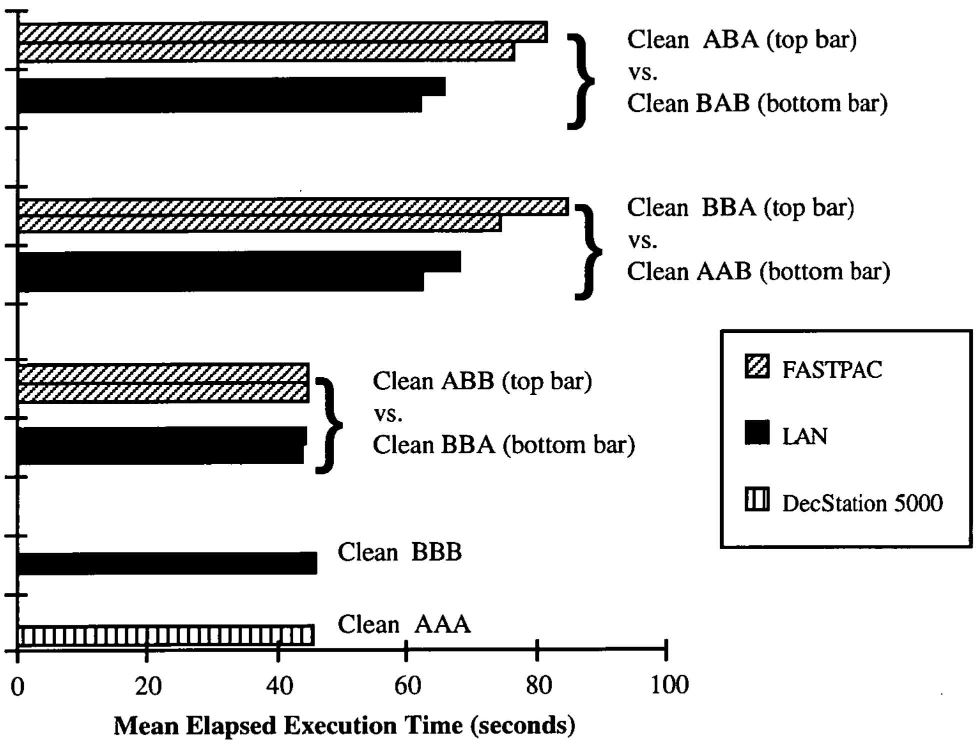


Figure 5.15

Execution Times of Corresponding Usage Configurations *ARC/INFO CLEAN* Command

Note: "CLEAN (*X-Y-Z*)" indicates an operation where required input files were read from disk at Workstation *X* and processed at the User's Workstation *Y*. Output files were then written to disk on Workstation *Z*.

5.2.3 INFLUENCE OF NFS OPERATIONS AND SETTINGS

A clear trend was observed in both the 1991 and 1992 experiments in any cases where the operations involved data stored on remote NFS-mounted directories. As mentioned in Sections 5.1.3, 5.2.1 and 5.2.2, any operations which employed NFS to transparently *write data to* a remote disk took substantially longer than corresponding operations which used it to *read data from* the same disk. Other recent GIS research efforts also suggest this phenomenon was being caused by the inherent nature of NFS itself [Hammer, 1992]. While no quantitative estimates or results were included, Stern [1991], Katz [1991] and Bachmann et al. [1989] also mention the effects of this behaviour on file transfer and copying operations.

In order to understand the causes of this phenomenon, some background is necessary. Introduced in Section 2.4.3, the Network File System (NFS) allows users to transparently access and use remote disks in the same manner as those on their own machine. Because of this, implementations of UNIX NFS and PC-NFS have become extremely popular and common in many organisations [Katz, 1991].

For reasons of data consistency and transmission security during an NFS *write* operation, packets must be sent to the remote workstation, written from memory to disk, and an acknowledgement packet sent back before the next shipment can be transmitted across the network. Waiting for these acknowledgements leaves both the sending disk and the network itself underutilised, thereby slowing everything down. NFS has been designed such that writing data <u>to</u> a remote disk takes substantially longer (2-3 times) than reading it <u>from</u> there.

This has a significant impact on performance within the LAN and appears to have a disproportionate effect on performance across FASTPAC. For example, *CLEAN* and *COPY* operations which write the output to a remote disk (i.e., those operations labelled AAB or ABA) took *consistently* longer than corresponding operations which read input data from remote disks but stored the output locally. (See Figures 5.8, 5.9, 5.13 and 5.14.) Due in part to the much higher number of acknowledgement packets involved, corresponding increases in execution times across FASTPAC were markedly higher as well.

NFS tuning itself also appears to be an important consideration in applications performance across both the LAN and the FASTPAC cloud. Research indicated that the number of "NFSD" daemons[1] specified when setting up a user's session control parameters can have a significant influence on performance — particularly when dealing with large image files. Figure 5.16 compares FASTPAC / LAN execution times for selected GIS operations when: (a) eight (8) NFSD daemons are specified at start-up (the number recommended by Sun); and (b) zero (0) NFSD daemons are running.

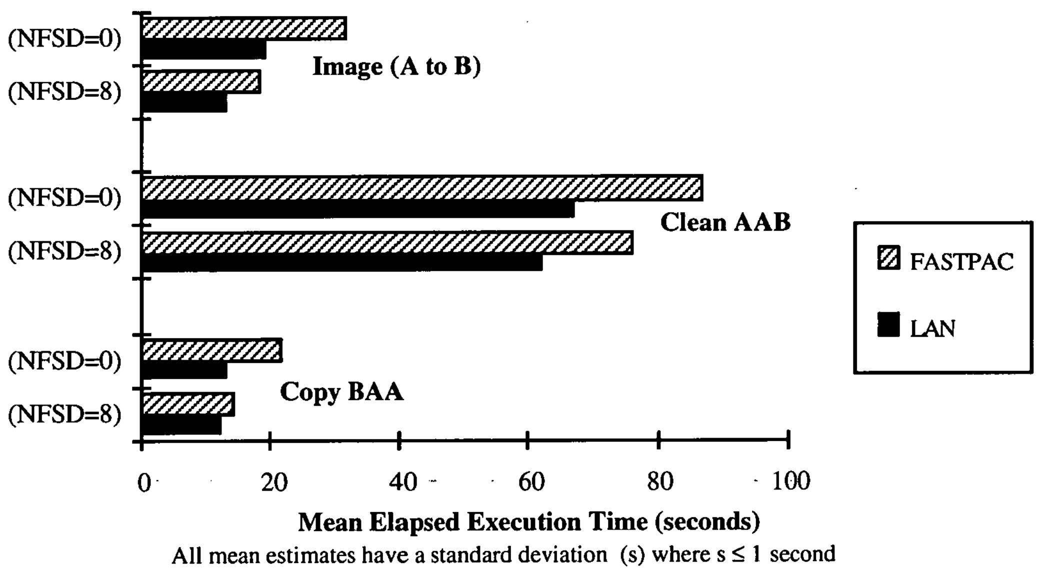


Figure 5.16
Influence of the Number of NFSD Daemons on Performance

---

[1] A "daemon" is a small, special-purpose computer program which runs continuously in the background and is designed to complete a specified task when a particular operating-system level command is invoked (e.g., calling to a printer or a remote disk).

NFS-related daemon setup work is often regarded as a "black art" and any variations to system defaults are usually left to system and/or network managers at a customer site [Stern, 1991]. In practice, management must ensure that such people appreciate that NFS settings and performance characteristics can have a noticeable and sometimes disproportionate influence on system response times across FASTPAC.

The adverse effects of NFS on overall performance are being addressed by a number of vendors who are now offering add-on accelerators for workstations. For example, Sun Microsystems now markets a "Prestoserve" board [Sun Microsystems, 1991] which increases the speed of NFS *write* operations and significantly improves performance across a LAN under normal operating conditions. Products like Prestoserve may reduce incremental differences between LAN and FASTPAC times as well, but further research would be required to verify this.

5.3 GIS Performance Under Varying Network Traffic Conditions

The tests completed on dedicated networks are useful in providing baseline figures for the performance of various GIS operations in different configurations. However, most "real-world" users must compete for the resources of both the network and the file server(s) in their day-today operations. This section describes the results of four separate sets of experiments which examined the effects of varying server and network traffic loads on GIS performance in a client-server environment.

5.3.1 PERFORMANCE UNDER MULTI-USER CONDITIONS

Background

The test results described up to now have dealt with dedicated servers and network bandwidth. Under these conditions, LAN Analyzer observations indicated that only 3-5% of the network bandwidth was used on average for GIS-related operations, with maximum bandwidth utilisation never exceeding 22%.

While these experiments did provide useful baseline results, such conditions rarely exist in practice. In an attempt to obtain more "realistic" results, separate tests were undertaken during the 1991 experiments to see how response-time performance was affected by other GIS users on the network. In a limited series of tests, the respective response times required to complete selected GIS operations in a given script were measured and compared under conditions where: (a) users were working alone; (b) two users were working simultaneously and accessing the same server; and (c) two users were working simultaneously, but accessing different servers.

Selected results from this test are summarised in Table 5.1 and illustrated in Figures 5.17 and 5.18.

Results

The response times from these experiments exhibited much less consistency than earlier observations, with the standard deviations ranging from ±2 to 28 seconds. Absolute spreads were higher on much longer operations like UNION but, in relative terms, were generally less than 10% of the estimated mean times for the operation. The spreads in individual sets of observations were neither higher nor lower across FASTPAC than within a single LAN, indicating that any variations introduced by the network were small in comparison with those introduced by competition for disk resources. This is confirmed by noting that the standard deviations are much larger for those cases where two users are competing for access to the same disk.

The reasons for the higher standard deviations in this set of experiments stem from more than just the increased demand on network and server resources. Although the respective command scripts for these tests were ordered in the same manner and were started at the same time, the resource sharing and collision detection processes within the various hardware and network components introduced a time lag between the two sets of processes. As a result, corresponding operations eventually became out of phase with one another as the iterations mounted. Since the competition for network and disk resources varies from iteration to iteration depending on the corresponding commands being executed by each script at a particular time, standard deviations tend to be much higher for these tests.

TABLE 5.1: SUMMARY OF SELECTED RESULTS PERFORMANCE UNDER MULTI-USER CONDITIONS

		RESPONSE TIME (secs)		STANDARD DEVIATION	
		LAN	FASTPAC	LAN	FASTPAC
CLEAN	clean (A-B-A)	67.80	78.30	1.87	1.49
	clean (A-B-A)**	67.00	72.60		2.76
	clean (A-B-A)*	88.22	102.70	3.38	7.17
	clean (A-C-A)	73.07	93.20	1.16	0.63
	clean (A-C-A)*	94.78	117.80	3.87	15.30
	clean (B-C-B)	73.80	81.10	1.15	0.57
	clean (B-C-B)**	81.30	96.30	9.94	5.60
COPY	copy (A-B-B)	24.30	27.30	2.50	0.82
	copy (A-B-B)**	27.20	29.90	4.76	1.37
	copy (B-C-A)	47.20	51.40	1.01	0.70
	copy (B-C-A)**	53.60	55.20	13.71	5.18
	copy (A-B-A)	47.80	53.00	0.79	1.41
	copy (A-B-A)*	96.00	90.50	16.48	15.13
	copy (A-C-A)	51.67	58.70	0.82	1.16
	copy (A-C-A)*	97.90	91.70	13.30	14.15
DRAW	draw (A to B)	32.60	34.20	0.52	0.42
	draw (A to B)**	34.90	35.50	2.96	1.58
	draw (A to B)*	35.20	35.80	1.48	0.63
	draw (A to C)	34.67	39.00	0.62	0.94
	draw (A to C)*	39.50	45.10	2.95	2.85
	draw (B to C)	35.33	39.00	0.72	0.47
	draw (B to C)**	34.90	40.60	1.66	3.06
UNION	union (A-B-B)	315.60	324.20	1.84	6.81
	union (A-B-B)**	330.30	342.10	12.40	11.73
	union (A-B-A)	398.60	446.30	1.35	4.03
	union (A-B-A)*	619.30	592.90	13.40	19.86
	union (A-C-A)	414.80	475.00	2.34	1.94
	union (A-C-A)*	611.20	600.30	8.94	50.27
	union (B-C-B)	436.07	455.70	3.37	1.70
	union (B-C-B)**	440.70	497.30	28.51	24.74
	LEGEND				
	(1) Command (X-Y-Z) means that the input data is read from the disk on machine "X" to be processed on Machine "Y", then written to disk on machine "Z".				
	(2) Command (X-Y-Z)** — performed when another user is accessing the same disk at the same time.				
	(3) Command (X-Y-Z)* — performed when this and another user are sharing two disk drives located in different places across the network.				

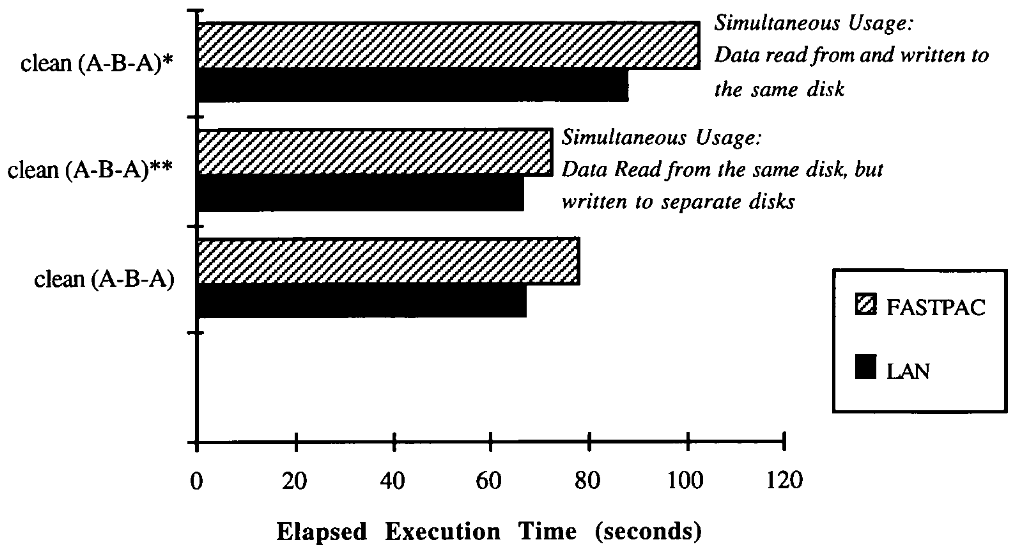


Figure 5.17
Comparative Execution Times for Concurrent Usage
ARC/INFO CLEAN Command

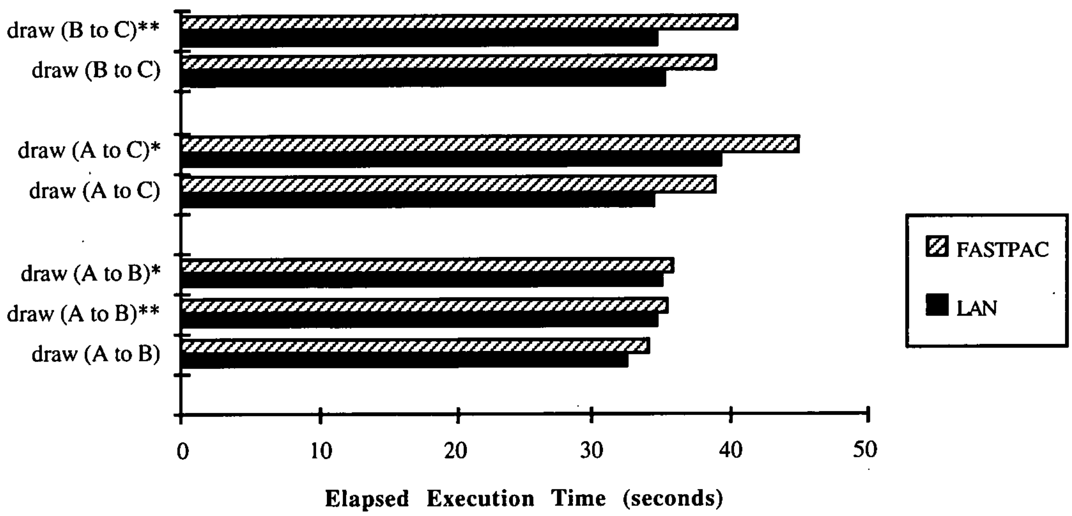


Figure 5.18
Comparative Execution Times for Concurrent Usage
ARCEDIT DRAW Command

The response times listed in Table 5.1 illustrate how performance varies under different conditions. The worst-case scenario (those marked with a single asterisk *) is that where a user from a diskless client is competing with a second user for the same server's NFS and disk resources at roughly the same time. As a result, the configuration becomes disk-bound and performance times increase substantially. Despite this direct competition for bandwidth and disk resources, however, no operation consistently takes twice the time to complete. In many cases, the times required for display-type operations are only increased by a few seconds.

The second set of tests (marked with a double asterisk **) involves a modified situation where input data files have been distributed between two servers (one on each side of the network, in the FASTPAC configuration). In certain cases, the times required were actually *shorter* than those encountered under single-user conditions since the two disks in question were apparently being utilised concurrently for disk *read* and *write* operations. This was not true in all cases, however, and performance will vary depending on the data handling techniques employed by the software designers in each operation. Even so, distributing data across different servers on the network appears to improve performance sufficiently enough to make it cost-effective to complete even processing-intensive operations concurrently.

Finally, while not true in all cases, the differences were usually larger in cases where the command was invoked from the diskless workstation C. Once again, these longer times can be traced primarily to the increased paging demands involved in swapping data to and from memory onto a remote disk located either elsewhere on the same LAN or across the FASTPAC cloud. In this case, however, the differences have been exacerbated by greater competition for network and server resources.

Assessment

These tests represented an important first step in examining the response-time performance of a variety of GIS operations in a client/server environment under multi-user conditions. Preliminary results indicated that the network and disk resources can all handle higher levels of GIS traffic on the network before filling up. Comparing the respective performance of the different configurations further suggested that, if data and resulting disk usage can be allocated strategically among different servers in a network, performance can approach optimal levels even with heavy GIS traffic loads. This supports observations by Dowers et al. [1991] in their research and supports the practical experiences of network managers working with large GIS data files in a client/server environment [Zhou, 1991].

In summary, then, these tests illustrated how response-time performance can vary depending on the overall load on the network and disk resources. However, since third party communication services (like Telecom Australia's FASTPAC) will often carry other traffic besides just that of the GIS users, it is important to determine the respective influences of *network* vs. *server* loading on overall GIS performance. Experiments towards this goal will be discussed in the next two sections.

5.3.2 VARYING THE NETWORK TRAFFIC LOADS — 1991 EXPERIMENTS

Background

To control the levels of background data traffic on the network, a Hewlett-Packard LAN Analyzer unit was placed on the same side of the FASTPAC cloud as the Server (A) and programmed to broadcast specified levels of packet traffic to an unknown destination address (i.e., across both LANs and the FASTPAC cloud). A subset of the GIS operations were performed in a configuration where the user operated from a diskless workstation (C) and both the software application and the data resided on the Server.

Table 5.2 contains the specific network configurations and simulated load conditions under which GIS execution times were measured.

Table 5.2: Network Loads and Configurations Used in Testing

Network Load Condition	LAN	FASTPAC
0% Background Load; Dedicated Network	✔	✔
Synthetically-generated data traffic simulating 5% Utilisation of available bandwidth; Two (2) 60-byte packets transmitted back-to-back	✔	✔
17 % Utilisation of available network bandwidth; Two (2) 60-byte packets transmitted back-to-back	✔	✔
20% Utilisation of available network bandwidth; Two (2) 60-byte packets transmitted back-to-back	✔	
17% Utilisation of available network bandwidth; Packets of random size (60 - 1514 bytes) transmitted at random intervals	✔	

Tables 5.3 and 5.4 summarise the results of the first series of tests monitoring GIS performance under controlled network traffic loads (completed in 1991). Due to time constraints during the testing period, each set of operations was only carried out once. Therefore, the figures in this table should be used to indicate *level-of-magnitude differences* only.

Once again, it must be understood that these early experiments were designed to test FASTPAC under maximum loading conditions. A substantial number of users — all involved in file transfer or display activities — would be required to generate and maintain a constant 10% network utilisation level. In light of the much lower levels observed in practical network monitoring experiments (as discussed in Chapter 3), such sustained traffic loads would seldom occur in practice.

Table 5.3: **Execution Times Required for GIS Commands Under Differing Broadcast Traffic Loads**
LAN Configuration

Nature & Level of Background Load / GIS Operation	Original Values 0% Background Traffic	5% 60-byte packets in constant stream	17% 60-byte packets in constant stream	20% 60-byte packets in constant stream	17% Random-size packets, randomly-generated
draw (A to C)	34.67 secs	35 secs	35 secs	35 secs	40 secs
copy (A-C-A)	51.67	52	54	53	59
union (A-C-A)	414.8	415	415	419	475
clean (A-C-A)	73.07	74	82	75	95
copy (A-C-B)	49.27	49	50	49	51
reselect	1.64	1	1	1	2
polygons (A to C)	4	5	5	5	5
polygonshades (A to C)	6.4	7	8	6	7

Table 5.4: **Execution Times Required for GIS Commands Under Differing Broadcast Traffic Loads**
FASTPAC Long (FP[B]) Configuration

Nature & Level of Background Load / GIS Operation	Original Values 0% Background Traffic	5% 60-byte packets in constant stream	17% 60-byte packets in constant stream
draw (A to C)	39 secs	39 secs	87 secs
copy (A-C-A)	58.7	59	181
union (A-C-A)	475	473	907
clean (A-C-A)	93.2	95	360
copy (A-C-B)	50.9	49	94
reselect	2	2	5
polygons (A to C)	4.8	5	14
polygonshades (A to C)	6.9	6	9

General Observations

While results from these preliminary tests cannot be considered conclusive, they did further the belief that GIS-related performance across FASTPAC is sensitive to both the volume and composition of background broadcast traffic. Key observations arising from these tests include:

(1) There appeared to be little difference in the execution times observed in any of the configurations when a 5% constant traffic load was placed on the network.

(2) In the LAN configuration, there was also little apparent difference in the execution times observed when constant loads of either 17% or even 20% were placed on the network.

(3) Overall response times on all operations increased slightly when the nature of the loading was changed from a constant stream of small packets to the randomly-generated stream of random-sized packets. In retrospect, this appears to reflect the influence of large background packets on Ethernet LAN performance rather than on traffic through the FASTPAC cloud.

(4) While little difference in performance was observed at loads of 5% and 15% , execution times in the FASTPAC configuration did increase when a 17% background load was applied. This is understandable since — in this particular case — the broadcast background traffic was generating additional packets across both LANs and on the FASTPAC cloud.

Even in these cases, however, the observed delays should not adversely affect usage of the system: display times were still reasonable for interactive usage and -- since batch processing jobs would often be completed in the background or in off-hours -- the level of increases seen here may not be important to many users.

Note: In contrast to the other operations which took longer as loading increased to 17%, there was no noticeable time increase observed for the "COPY A-C-B" operation. In this case, the user read data from the Server on a different LAN across the cloud but then wrote it to another disk *on the same LAN*. This supports previous observations which suggest that operations which *write to-* (rather than *read from-*) remote disks take longer to perform.

(5) Delays across the LANs and the FASTPAC cloud increased substantially when background traffic levels were increased to 17%. Even the COPY operation which wrote data to another disk on the same LAN took much longer to complete under this degree of network loading. Under *these* circumstances -- while lower-priority, processing-intensive operations might still be run unnoticed in the background and deliver acceptable throughput -- the increased execution times for display-type operations *would* be noticeable.

(6) We attempted to take measurements at 20% utilisation under the FASTPAC configurations, but the additional packet traffic present on both LANs and across the FASTPAC cloud led to the client workstations being unable to even keep the UNIX operating system running successfully.

Assessment

These tests provided a preliminary indication of how GIS response times changed as network traffic loads varied across the network. However, while useful, subsequent examination and analysis identified two important flaws which would have to be addressed in the next round of experiments. Specifically:

(1) To obtain a more complete and reliable picture of performance, repeated sets of observations were required. Given the higher spread expected to be present in

high-traffic situations, the observations would have to be repeated at least 20 times under the same controlled conditions as discussed in Chapter 4.

(2) Discussions with network analysis specialists from Telecom Australia suggested that — by broadcasting to a *random destination* — the random traffic settings employed on the LAN Analyzer in these experiments actually generated much higher levels of network traffic than originally specified. Rather than simply broadcasting packets to all destinations, it was decided that the traffic generator settings should be respecified to only send packets to (depending on the test): (a) one specified workstation on the same LAN or (b) one specified workstation on the other side of the FASTPAC cloud.

These deficiencies were addressed in the 1992 round of experiments conducted in Melbourne, and the respective influences of *both* network and server loads were examined at that time. The results of those later experiments are discussed in the next section.

5.3.3 VARYING BOTH THE NETWORK AND SERVER LOADS — 1992 EXPERIMENTS

Description

As with the 1991 experiments, GIS performance was tested under sustained traffic loads of varying degree and composition using the "Traffic Generator" program on the LAN Analyzer. It was originally intended to generate background traffic patterns similar to those observed on an actual operating network (as discussed in Chapter 3). However, since project scheduling constraints prevented this, much simpler messages were eventually coded and used in the experiments. Varying levels of background traffic possessing three different compositions of messages were generated and a subset of the performance tests was run once again.

In other tests, the effects of variations in *server* loading on the response times were also examined. In these experiments, the GIS operations involving a remote server were invoked while that server was in the process of handling other (locally-generated) GIS operations. Results were obtained for experiments involving zero, medium and heavy server loads.

A full summary of the actual results is contained in Appendix C.4.

Results of Network Load Testing

Table 5.5 summarises the measurements made during this testing. The sharpest declines in performance across FASTPAC were observed under sustained loads composed of small (60 byte) packets. While this is generally true for LANs as well, performance across FASTPAC declined to the point where the application could not run across the connection at sustained loads of over 10%. In practice, no network would operate with a sustained load of such small-packet traffic unless it served a very large number of busy users all working on interactive text terminals at the same time, and FASTPAC is not aimed at such a market.

FASTPAC *was* able to sustain higher levels of traffic containing large packets (as would be found during heavy file transfer activity) or a mixture of large and small packets (as found on a busy NFS network). However — under the varying loads — GIS performance did decrease at different rates for different operations, with the operations which involved *reading* data from a remote server being more adversely affected. (See Figure 5.19) While the time required to complete operations involving *writing to* a remote disk also increased, the relative increases were generally much smaller.

It is possible that these uneven effects on performance may be due to the same NFS-induced delays described earlier in Section 5.2.3. The effects of higher network load on NFS *read* operations —which can draw large streams of data across the network quickly — should be proportionately much higher than on *write* operations which send data in smaller clusters and wait for an acknowledgement from the remote machine before sending the next shipment. *Read*-remote and *Write*-local operations still take much less time to complete overall (as shown by the times for *COPY ABB*), but the proportionate effect of network loading on their performance is much higher.

Assessment

Even at the highest sustained utilisation levels tested, the mean response times did not increase by more than 40% over those observed when no other network traffic was present.

In fairness, these findings should not imply that network loads — and therefore response-time increases — will *never* exceed this amount. Quite the contrary, in the network monitoring experiments cited by (e.g.) [Fowler et al., 1991], peak LAN utilisation levels exceeded 50% for very limited periods of time during large file transfers, back-up operations or image displays from data stored on a remote NFS-mounted disk. Under such load levels, one should expect to experience much longer delays in the completion of certain GIS operations.

Clearly, heavier sustained GIS traffic loads may be found on LANs containing a larger number of workstations and servers than those found on the Victoria Conservation and Environment networks [Healey, 1994]. Practically speaking, however, it would be uncommon to expect *sustained* LAN utilization levels of 30% or more on a well-designed operating network. Since many sites usually restrict heavy data backup activities to off-hours, these spikes on a LAN should usually be of short duration during the normal working day.

Table 5.5

GIS Performance Times Under Varying Background Traffic Levels

(All Performance Times Shown in Seconds.)

	MESSAGE SIZES							
	Zero	Gr. 1	Group 2			Group 3		
% Background Traffic	0%	5%	10%	20%	30%	10%	20%	30%
Copy BBA	58	62.6	58.4	59.2	61.2	60.8	62.8	65.6
Copy ABB	22.2	26.8	24.8	26.2	28.2	24.4	26.6	30.2
Copy ABA	73	83	77.5	76.6	85.6	77.8	81.4	87.8
Image (from A to B)	31.2	38.2	35.2	37.2	40.8	35	39	43

Note: Copy "XYZ" indicates an operation where input files were read from disk at Workstation *X* into memory on the User's Workstation *Y* and then copied onto the disk at Workstation *Z*.

Message Sizes Legend

Zero — Dedicated network; No background noise levels

Gr. 1 — Three (3) 60-byte packets; Three (3) frames per burst
— Note: This message composition put significant strain on FASTPAC. The GIS application could not be run at 10% load.

Group 2 — Three (3) 1514-byte packets; Three (3) frames per burst

Group 3 — One (1) 60, One (1) 256 and Two (2) 1514 byte packets; Three (3) frames per burst

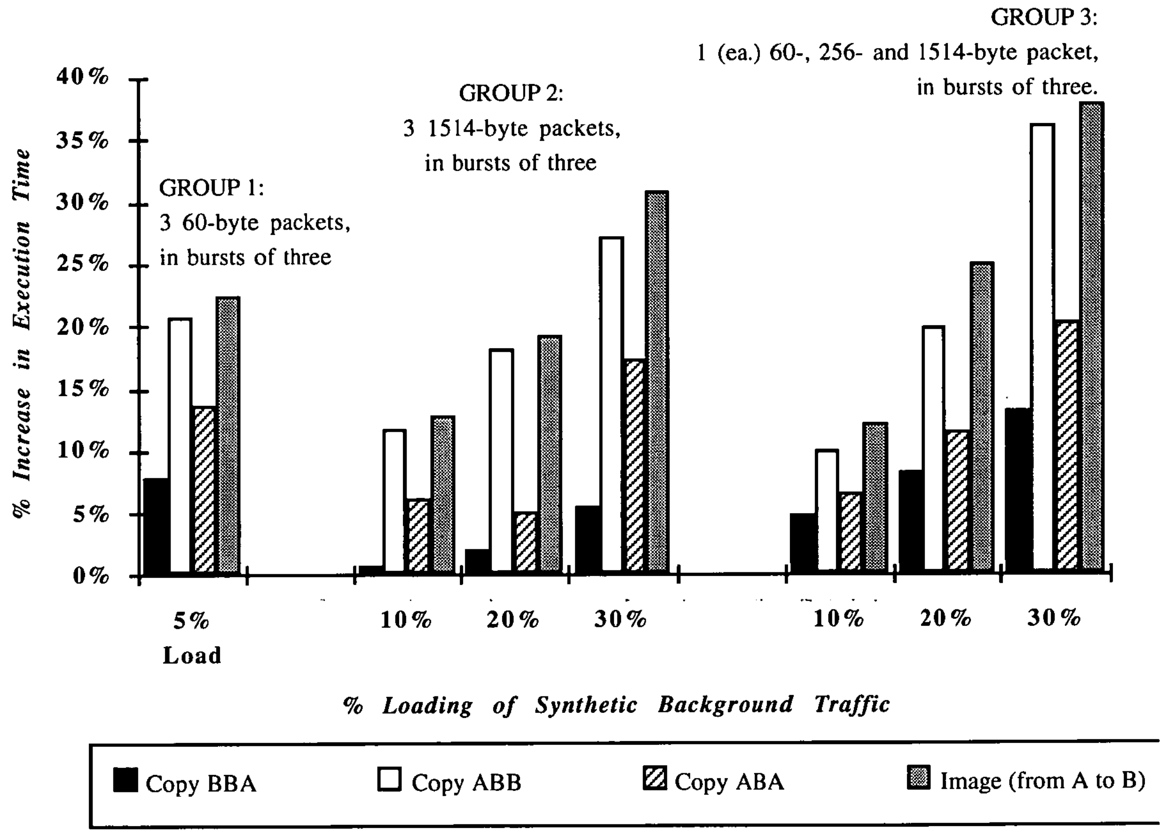


Figure 5.19

Influence of Background Traffic across FASTPAC Network

Relative Increases in Execution Time for Selected GIS Operations

Results of Server Load Testing

Description

The purpose of the next series of experiments was to isolate the effects of additional loading on the server *alone* on GIS response-time performance. To avoid increasing overall traffic on the network as well, separate AML scripts were developed and then executed on the Server (A) itself. Two different scripts were designed to execute a series of *ARC/INFO* commands which would place (respectively) a variable and a heavy load on the server. Measurements made during this testing are summarized in Table 5.6 and illustrated in Figure 5.20.

The script used to generate a variable server load consisted of an *ARCEDIT* DRAW command (to draw a 750 Kbyte file), a second DRAW command to display the same file once again while the data was still in memory (thus minimising the disk usage for that period) and an *ARC/INFO* COPY command (to duplicate a 3.5 Mbyte coverage). By comparison, the AML script used to generate a heavy server load consisted of three separate ArcInfo COPY commands which duplicated three different ArcInfo coverages.

In both cases, the order of commands in each script was designed to minimise any chance of memory cacheing at the workstation (i.e., holding a given coverage in memory between operations, thereby reducing the disk accesses required and distorting the resulting response times).

While each of these scripts was being run on the server, a separate AML script was being executed on a remote workstation. This script contained a series of CLEAN, COPY, DRAW and IMAGE operations which read and/or wrote data from that same server. The usual logging and performance timing tools were then used to record response times for each of the operations invoked.

Table 5.6: Influence of Additional Workload on Single Server
Response Times for Selected GIS Operations
Under Zero, Variable and Heavy Server Loads

	MEAN VALUES (seconds)			*STANDARD DEVIATIONS (seconds)*		
Operation	Zero	Variable	Heavy	Zero	Variable	Heavy
Clean ABA	68.40	87.40	131.50	1.17	11.01	3.54
Clean BBA	66.90	85.60	125.00	1.03	8.22	N/A
Copy ABA	49.10	68.60	120.00	0.32	14.10	1.41
Copy ABB	13.50	16.95	24.00	0.53	3.00	N/A
Copy BBA	39.90	55.55	95.50	0.32	7.88	2.12
Draw (to B from A)	29.30	30.30	32.00	0.67	1.17	N/A
Image (to B from A)	19.00	24.25	44.00	0.00	6.65	1.41

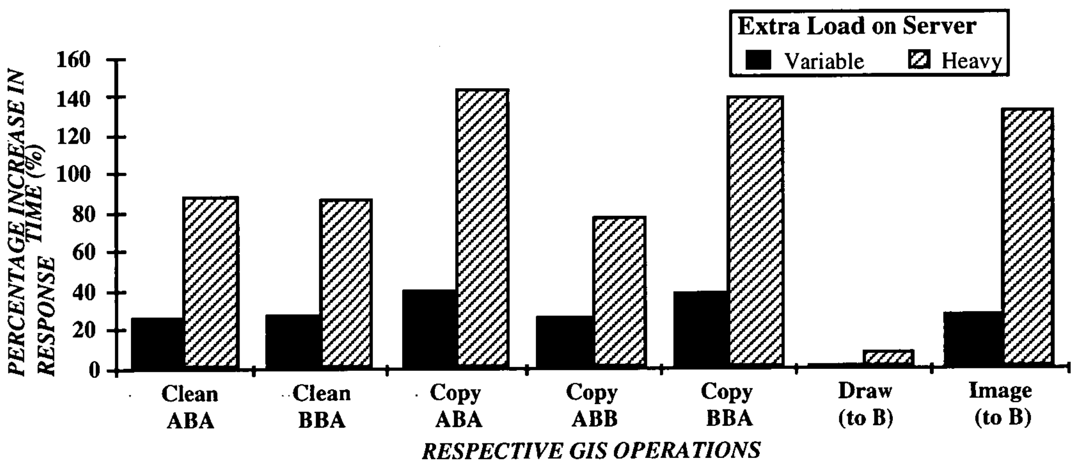


Figure 5.20

Influence of Additional Workload on Single Server
Relative Increases in Execution Time for Selected GIS Operations

Results

Table 5.6 and Figure 5.20 clearly indicate the significant effect of server loading on the response-time performance of selected operations. Even the "variable" load slows most operations down by at least 25%, with the "heavy" load increasing most times by at least 85% and some by up to 145%. The only exceptions to this involve the *DRAW* operations, which only increase by 3% and 9% respectively. These lower increases may be due to the fact that these particular operations deal with a much smaller file than those required for the other commands.

Further, it appears that those *COPY* operations which involve writing data *to* the server take proportionally much longer than corresponding *CLEAN* operations. This is probably due to the fact that the *CLEAN* operation has a higher processing component which is unaffected by additional loading on the server.

Finally, it is interesting to note that the respective behaviours of the *COPY* operations carried out in this series of experiments are different from those carried out in the previous network loading tests. This may be due to the fact that contention-handling techniques on the Ethernet (which would affect the previous tests) differ from those on the disk, which would be controlled by a combination of hardware and software.

Assessment

These experiments indicate that, at least for the configuration tested, moderate levels of server activity exert at least as much influence on GIS application performance as sustained network traffic loads. Heavy and sustained server loads can result in even greater and more pronounced performance delays. Understandably, I/O-intensive operations like CLEAN and COPY take considerably longer to perform under variable and heavy loading conditions. At least with the current generation of disk technology, this finding reinforces the strategy of distributing data over a number of different servers in the network to optimise performance.

5.4 Discussion

Experiments outlined in this chapter have attempted to quantify GIS response-time behaviour in a client/server environment under varying network and server loading conditions, and also identify the factors which may affect performance in each case. This section briefly summarises the results once again and discusses the implications of performance delays with respect to end-user satisfaction.

5.4.1 SUMMARY OF RESULTS

(1) **There was little difference between FASTPAC and LAN configurations in any of the tests involving UNIX *ftp* transfer of large image files between workstations or UNIX *cp* copying of GIS files between NFS-mounted directories across the network.**

Test results indicated that — for *"ftp"* file transfers between workstations — the execution time differences between LAN and FASTPAC configurations were almost negligible. More noticeable time differences were observed in certain cases when using the *"cp"* command to copy files between LANs across the FASTPAC link, but these apparent delays were due to Network File System limitations rather than any characteristic of FASTPAC.

(2) **While GIS application performance across FASTPAC was measurably slower than that observed across the LAN, most execution times were still within reasonable limits.**

Test results indicated that the FASTPAC link added measurable delays to many of the operations tested. However, the execution times of all selected

GIS display and processing operations on <u>unloaded</u> networks increased *at most* by 40% when run across FASTPAC (usually much less, depending on the configuration), and the observed values were generally quite consistent within each sample. In many cases, the time differences between stand-alone and LAN configurations were larger than any incremental delays introduced by the FASTPAC link.

(3) **In all client/server configurations tested, GIS response-time performance for a given operation varied significantly depending on the respective physical locations of the input and output data files specified by the user.**

Due to the nature of NFS itself (discussed in Section 2.4.3), *writing* data to a remote NFS-mounted disk takes substantially longer than *reading* it from one. Further, due to the number of acknowledgement packets which must be sent back and forth through the FASTPAC cloud, this same NFS-*write* characteristic appears to disproportionately increase performance delays across FASTPAC.

This will affect any operation which deals with files stored on a remote disk. While exceptions were observed, the performance of the various possible configurations can generally be ranked in the following order:

Input read from:	***Output written to:***	***Ranking***
Local Disk	Local Disk	1 (Fastest)
Remote Disk	Local Disk	2
Local Disk	Remote Disk	3
Remote Disk	Remote Disk	4 (Slowest)

Given the performance differences observed, it may still be most appropriate to transfer copies of these files to the host workstation (or at least the local server), perform the necessary operations, then transfer back all or part of the modified files on completion. This would improve overall turnaround time on processing-intensive operations and would minimise the amount of data and NFS traffic across the FASTPAC cloud. However, it would require a stand-alone workstation on the desk of every individual using the system for more than just routine query and display — an expensive proposition even when hardware prices are falling.

X-terminals are considerably less expensive than workstations and work well in applications which primarily involve data display. However, the demand on server resources may prove problematic if large numbers of users routinely process large GIS files during normal working hours. Diskless workstations, while potentially reducing the load on a central server, would substantially increase levels of data traffic across the network due to paging operations and data transfer demands.

Dowers et al. [1990] suggest that *dataless* workstations (i.e., those with enough disk space to minimise paging across the network, but which still must store the resulting files on a remote disk) may represent an ideal compromise between stand-alone and diskless workstations. However, that particular research was undertaken over a network running DECNet rather than NFS. Configurations which store no data files locally (corresponding to the fourth row of the previous table) are the most affected by NFS characteristics and may offer the slowest performance of all configurations in NFS-based networks.

Understandably, throughput is only one consideration in a corporate data management strategy and must often be traded off against data security and logistics considerations as well as capital versus operating cost tradeoffs. For example, if customer organisations interconnect their LANs using third-party services like FASTPAC, incremental data transmission costs may quickly accumulate if large files are constantly being transferred (or NFS-written) to remote disks across the FASTPAC cloud by large groups of users. Such cost tradeoffs will be discussed further in Chapter 6.

(4) **FASTPAC configurations delivered satisfactory (and competitive) performance on GIS display operations in configurations involving diskfull workstations.**

Any operations involving writing data across the network experienced some slowdowns in performance when using FASTPAC linkages. However, performance-time degradations when invoking graphics *display*-type operations— particularly when using disk<u>full</u> workstations — were minimal (e.g., < 2 seconds, or ~5% of overall speeds). The final section of this chapter will discuss the tradeoffs involved in deciding what may or may not constitute a "satisfactory" response time. Although the response times observed on IMAGE operations across FASTPAC were up to 40% longer, the absolute differences were still judged to be within satisfactory limits (i.e., <10 seconds).

This is an important observation, since — once databases are initially loaded — the number and proportion of query and display-type operations on the network will grow substantially as new end-users come on-line. This phenomenon is evident from the investigation of GIS usage in participating organisations described in Chapter 3. Rather than any special analyses or complex enquiries, most of the routine GIS-related activity attributed to end-users involved image retrieval and display of standard datasets. Other technical or operational factors notwithstanding, FASTPAC linkages may provide a degree of performance to such users close to that of a LAN *regardless of whether the data is stored locally or at a remote site.*

(5) When working across the LAN or FASTPAC, the physical location of the GIS application software appears to have only a minor effect on overall performance.

Comparisons of similar usage configurations indicated that the execution times to complete the comparable GIS operations tested usually remained similar *regardless of whether the application was stored locally or resided on the remote server* on the other side of the FASTPAC cloud. Differences in some corresponding mean values *were* determined to be statistically significant at the ".01" confidence level. Practically speaking, however, none of the corresponding times tested differed by more than 20% of the overall time involved and most differed by less than 8%.

(6) The performance of the diskless workstations was markedly slower than other machines in most applications.

The performance of the disk<u>less</u> client workstation employed was observed to be 6-20% slower than its counterparts in handling any processing-intensive GIS operations in all three network configurations. The most significant degradations were observed in I/O- and processing-intensive operations, where significant amounts of memory-to-disk-to memory swapping may be taking place and data had to be posted from the diskless client across the FASTPAC cloud to the remote server. By comparison, there were few significant performance differences between diskfull and diskless clients in any of the configurations on the display of graphics files.

Only 12 Mbytes of memory was installed on the diskless workstation at the time of the 1991 tests and, in fairness, performance may have improved by increasing the memory capacity of the machine. However, equipment and scheduling limitations and the demands of other tests prevented this factor from being tested further at the time.

(7) **The X-terminal equipment tested provided comparable response to normal workstations on all operations tested, but users must be aware of the network and server tradeoffs involved.**

In the 1992 experiments, since the Labtam X-Terminal was logged into one of the DEC workstations, the execution times for most processing and I/O-intensive operations mirrored those observed on the host machines themselves. (Indeed, packets containing the original command and the return prompt represented the only traffic across the net in these cases.) In display operations, while there was little difference in DRAW performance, it took substantially less time to complete the IMAGE command on the X-terminal. Rather than bringing across the entire file into local memory via NFS, the X-terminal only transferred the bit-mapped screen image itself.

Clearly, while the apparent performance is satisfactory in single-user situations, there is some concern over how well such systems would perform when many X-terminals were accessing the same server. The effects of server loading on GIS performance has already been observed and — if they invoke processing-intensive operations as part of their daily routine — X-terminal users would place a substantial extra load on the host server.

Dowers et al. [1990] suggest that, given the data volumes commonly encountered in spatial data handling operations, use of X-terminals is inappropriate in a GIS environment. However, this research indicates that X-terminals may not overload a server or a network if used primarily for data display operations. In any case, performance would ultimately depend on the nature, speed and capacity of the server in use.

(8) Performance figures for GIS tests run concurrently indicate that the network and disk resources can all handle higher levels of GIS traffic on the network before filling up.

Despite the direct competition for bandwidth and disk resources in those particular experiments, no operation consistently took twice the time to complete. In many cases, the times required for display-type operations were only increased by a few seconds. If disk usage is strategically allocated with this in mind, performance can approach optimal levels even with heavy GIS traffic loads.

(9) GIS performance was found to be sensitive to the network traffic load levels, but to varying degrees depending on the operation and the usage configuration involved.

GIS performance was tested under sustained traffic loads of varying degree and composition. The sharpest declines in performance across FASTPAC were observed under sustained loads composed of small (60 byte) packets. However, such sustained loads of such composition would usually be found on networks serving a very large number of interactive terminals all working at once, which would be highly unlikely.

FASTPAC was able to sustain higher levels of traffic containing large packets (as would be found during heavy file transfer activity) or a mixture of large and small packets (as found on a busy NFS network). GIS performance did diminish under the varying loads at different rates for different operations, with operations involving *reading* data from remote servers most adversely affected. However, as indicated by discussions in Chapters 3 and 5, *sustained* network utilisation levels of 30% or greater would seldom be tolerated in practice for very long across most operating networks.

Response times may vary depending on the levels of other "third-party" traffic on the same FASTPAC link, but modelling the effects of this are beyond the scope of this research. It is understood that other Telecom-funded research efforts are underway to ensure that real-time apportionment of FASTPAC bandwidth among different users and applications is handled "fairly".

(10) At least in the configurations tested, heavy server loading has a much more profound effect than background network traffic on GIS performance.

While figures will vary depending on the hardware in question, the experiments conducted here indicate that delays caused by heavy and sustained competition for disk resources will be larger than those caused by network traffic. While this is a problem in the short term, new server/storage technologies like disk arrays may minimise this particular bottleneck in the near future and place the performance onus back on the network itself [Katz, 1991].

5.4.2 "ACCEPTABLE" VERSUS "UNACCEPTABLE" RESPONSE TIMES

The results summarized throughout this chapter have dealt with response time primarily from a quantitative perspective. Absolute values have been given for response times under ideal circumstances, and a combination of absolute values and percentage-change figures have been used to describe the relative increases in response time across FASTPAC and under various loading conditions.

Are these incremental increases caused by inherent FASTPAC delays and/or competition for network resources significant *to the end user?* If so, all of them? Some of them? Which ones?

An examination of the human-computer interaction and human factors literature suggests there is no agreement over what constitutes an "acceptable" versus "unacceptable" delay. Studies on the effects of interactive response time variability are often conflicting and inconclusive [Myers, 1985]. Since people appear to be highly adaptive and will assume different strategies to cope with variations in performance, the various interpretations of "user satisfaction" may be speculative at best. One study [Geist et al., 1987] found that users' perceptions of response time variations were not accurate and appeared to be influenced by their most recent experience. Moreover, when these times would vary around a mean value, users tended to assume the fastest times as the "norm" and be frustrated with slower ones.

Given the current state of the technology in 1992, however, current thinking suggests that — especially in interactive or display-type operations — *consistency* in response time appears to be a key determinant of user satisfaction. For example, a recent study indicated that variations of less than ±50% of the average amount seem unlikely to significantly affect user performance [Mayhew, 1992]. Provided response times do not vary widely, users may be prepared to accept longer response times -- particularly in cases where they are not alternating between faster and slower systems If one accepts this 50% criteria as a measure of the allowable variance with respect to a mean value, then both the FASTPAC and LAN services appear to provide the desired consistency for most operations even under the heavy loading and multi-user conditions tested.

Finally, *predictability* of response is another key factor in ensuring user satisfaction. If users can understand the reasons for varying response times — and if these variations are predictable — they are more likely to accept them than when the response times fluctuate in a random and unpredictable manner. If users understand that response times in a client-server environment will be somewhat longer than on a stand-alone workstation — and that the degree of variation will depend on the number of users and types of activities also on the network at the time — system performance will seem more acceptable to them than if the sources of variability are invisible to them. Further, by appreciating these factors, they can predict periods during which response times may be better than others and plan their work accordingly.

Stretching Mayhew's findings to argue that a 50% overall increase in response times would still be acceptable may be justifiably questioned. In an environment where "fast is never fast enough", any tradeoffs involving longer response times will be opposed. However, from a practical standpoint, the acceptability of FASTPAC can be supported by at least three arguments:

(1) Unless users regularly move back and forth between a stand-alone and LAN-based workstations, they are not going to appreciate the differences in performance encountered in a client-server environment. *Consistency of performance* (i.e., low variation in response times) will probably be a more critical factor in such cases.

(2) In production environments, the processing- and I/O-intensive operations which demonstrated the highest absolute increases across FASTPAC (e.g., CLEAN and UNION) are currently regarded as time-consuming operations best handled in the background rather than as a front-line interactive task. Especially in cases where such operations are handled as a batch process after normal working hours (a common practice in many installations), the incremental increases possibly introduced by the FASTPAC cloud would seldom be noticed by most users.

(3) In cases where fast response times were important (i.e., image and graphics display operations), FASTPAC did perform satisfactorily. In the operations examined, response times across FASTPAC never increased by more than

40% over those measured on corresponding configurations on the LAN. Even in those cases (i.e., displaying a 6 Mbyte image), the maximum absolute increases involved were still under 10 seconds on lightly-loaded networks.

Given the growing number of GIS installations operating in a LAN environment — in the face of the often significant reductions in performance over stand-alone systems identified in these tests — it would appear that users are prepared to trade off the performance reductions in return for the greater data- and resource-sharing benefits found in a client/server environment. At least with the equipment and operations tested here, if users are willing to accept the performance degradations which occur anyway when moving from a stand-alone system to a LAN-based client-server environment, they should be able to accommodate the incremental differences introduced by the FASTPAC cloud.

While satisfactory response-time performance has been judged to be an important criterion in the acceptance and adoption of broadband communications technology, it is not the *only* criteria. Indeed, there are significant technical, operational and institutional concerns which must also be addressed if such networks are to be implemented in GIS organisations in the foreseeable future. After summarising the results of this research, the final chapters of this dissertation examine these factors and discuss the cost and management implications involved in implementing the FASTPAC technology in a GIS operation.

6.0 SUMMARY OF RESEARCH RESULTS

Controlled experiments were performed to assess GIS performance on selected operations across broadband metropolitan area networks using operations and network usage patterns often encountered in real-world sites. The results presented in Chapter 5 indicated that the FASTPAC service tested *could* deliver similar response-time performance to that of a local area network for many of the operations selected, although some operating system-related characteristics would have to be addressed in order to improve performance over the short-term.

These findings have important ramifications to both *individual users*, who may wish to use such networks only to draw large data files from a remote location, and large *programme-driven organisations* wishing to maintain centralised responsibility for data and system administration over a growing number of users in different locations. Before discussing the potential impact of broadband communications technology on GIS operations in the final chapter, it is worthwhile to briefly review the design and results of the experiments undertaken in terms of their contribution to proving or disproving the original hypothesis.

This chapter presents concise summaries of both the research methodology and results. Following this, conclusions are drawn regarding these results, the insights gained from the research and potential limitations or caveats which must be acknowledged. Finally, suggestions regarding potential follow-on research arising from these experiments are presented and discussed.

6.1 REVIEW OF HYPOTHESIS AND EXPERIMENT DESIGN

The original hypothesis leading to this research was:

"...broadband communications networks will provide the performance necessary to satisfactorily support the GIS application and data management requirements of a geographically-dispersed organisation from a single location."

To examine this hypothesis, a program of research was developed which attempted to examine the comparative performance of such networks in a manner which would provide relevant, reproducible and understandable comparative results to both end-users and managers. This research consisted of three integrated components:

(1) Selection of an Appropriate Metric for both Analysis, Comparison and Presentation of Results

The elapsed execution time (and differences in execution time) on a specified workstation platform was ultimately selected as the performance metric to be employed in this research. While its limitations as a metric are acknowledged, response time figures were still believed to be the ones best understood by end-users.

(2) Identification and/or Selection of Appropriate Test Parameters and Materials

- *Client-Server Usage Configurations:* To select network usage configurations which would be relevant to those found in actual use, an existing framework which examined the respective foci and trends in GIS usage under different organisational conditions was refined and extended for use in subsequent experiments. Following this, tests were designed which — using NFS-mounted directories and command syntax — varied the relative locations of the user, the input data and the output data on different workstations within the same LAN and between LANs across the FASTPAC connection.

- *GIS Operations:* To ensure the operations used in the testing were relevant to a significant proportion of "real-world" users, the author developed a defensible prototype approach to characterising and quantifying GIS command usage at a single site *and* for comparing usage between sites. This approach attempted to identify "representative" spatial data handling operations which would be either: (1) be commonly-invoked by users in various organisations; and (2) be I/O-intensive enough to "exercise" a network by generating heavy data traffic loads between client and server workstations.

- *Data Sets:* To ensure some degree of control and flexibility over the data in question yet still allow some measure of relevance to outside applications, a combination of "synthetic" and "real-world" data sets were employed in the testing.

- *Background Traffic Levels:* In an attempt to ensure the resulting range of execution times (and differences in time) for each operation were representative of those found in practice, the author used existing LAN analysis hardware and software to assess the nature and magnitude of GIS-related network traffic within and between different sites. These results were then compared with background traffic loads generated synthetically during the actual performance testing stages to ensure that the respective loads (and their resulting effects on performance) were representative of those which would actually be found under typical operating conditions.

(3) Examination and Analysis of the Comparative Performance of Selected Operations in a Client-Server Environment.

Employing the metrics, configurations, operations and datasets selected in previous stages, the author developed and tested a defensible, reproducible approach to determining comparative GIS performance levels in a client-server environment. Using this approach, the respective times required to complete selected operations on a stand-alone workstation and in client-server environments within a LAN and across FASTPAC were examined and compared on different hardware platforms (DEC and Sun workstations), under realistic resource allocation conditions, and with a mixture of workstations and X-terminals. Subsequent analysis examined the established or potential causes of major differences between respective sets of observations and suggested where delays might be significant enough to affect the original hypothesis.

6.2 DISCUSSION OF RESULTS

The research can be considered successful in the sense that — based on the criteria adopted for "satisfactory" performance [Mayhew, 1992] — the hypothesis *was* proven within the bounds of the tests performed: the incremental delays introduced by the FASTPAC service never exceeded 40% of the corresponding time measured across the LAN configuration. Since the testing schedule did not allow for a complete and comprehensive evaluation of performance for all commands, data sets, hardware/software configurations, etc., we cannot state with any certainty that the hypothesis is true in all cases. However, the increases in execution times observed across the FASTPAC service remained small enough to conclude that the system continued to deliver "satisfactory" performance on a number of GIS operations deemed to be representative of those found in actual working environments.

Research results demonstrated that — at least with the data files used here — the increase in execution time experienced moving from a stand-alone to LAN environment was usually found to be larger than the incremental increases observed when moving from LAN to WAN (i.e., FASTPAC). (See Table 6.1, the discussions in Chapter 5 and the tables in Appendix C.) This was considered to be significant (and unexpected) finding which reinforced the hypothesis: if users are prepared to accept the delays introduced when moving from a stand-alone to LAN environment, then the incremental delays introduced across the FASTPAC cloud should not unduly affect user satisfaction.

Table 6.1: Relative Increases in Execution Time Stand-Alone to LAN and LAN to FASTPAC

	MEAN EXECUTION TIME (secs)			*PERCENTAGE INCREASE*	
Operation	Stand-Alone	LAN	FASTPAC	*S/A to LAN*	*LAN to FP*
	(S/A)		(FP)		
Clean AAA	45.50				
Copy AAA	25.30				
Draw (to A from A)	28.70				
Image (to A from A)	7.40				
Clean BAA		44.40	45.17	-2%	2%
Clean ABB		44.47	45.00	-2%	1%
Clean BAB		62.33	76.60	37%	23%
Clean AAB		62.87	74.67	38%	19%
Clean ABA		66.00	81.80	45%	24%
Clean BBA		68.40	84.90	50%	24%
Copy BAA		12.60	14.50	-50%	15%
Copy ABB		12.93	15.20	-49%	18%
Copy BBA		37.93	47.80	50%	26%
Copy AAB		42.13	53.50	67%	27%
Copy ABA		51.07	61.40	102%	20%
Copy BAB		55.53	67.00	119%	21%
Draw (to A from B)		27.47	28.00	-4%	2%
Draw (to B from A)		28.87	29.60	1%	3%
Image (to B from A)		13.27	18.50	79%	39%
Image (to A from B)		13.33	18.50	80%	39%

Note: "*Command* XYZ" indicates an operation where required input files were read from disk at Workstation X, processed at Workstation Y, and the results written to disk on Workstation Z.

The research results also indicated that — due to different I/O demands in each case — some operations take proportionally much longer than others to complete across a network depending on the respective locations of the user, the input data files and the output data file. While the effects of the Network File System characteristics effects on GIS performance were raised in Hammer [1992]), this research identified and quantified the influence of NFS on performance of the same GIS operation under different usage configurations.

Finally, the original hypothesis was successfully confirmed under varying background network traffic loads. The research *did* demonstrate that GIS performance was affected by network loading, and the effects of a given load were higher across FASTPAC than across the local area network. However, even if one assumes that the network utilisation levels observed at the Victoria Conservation and Environment were low in comparison with other GIS user sites, the sustained loads required to *consistently* produce unacceptable performance across FASTPAC would be much higher than those which would probably be tolerated for long in actual practice.

While these experiments were designed to represent actual operating conditions in many respects, there are limitations which render the overall research effort only partially successful. For example, the absolute performance figures obtained must be considered applicable only to the specific hardware, software and network configurations tested the machines and O/S tested. Further, there is no evidence here to indicate that similar relative performance differences will exist between corresponding stand-alone and LAN-based operations performed on other equipment. Since some systems (Intergraph, for example), utilise the XNS protocol rather than TCP/IP to carry data across a LAN, the stand-alone vs. LAN vs. FASTPAC performance on those systems may differ significantly from that observed in these experiments.

Most significantly, while efforts were made to conduct the experiments under conditions "representative" of those found in practice, time constraints did not permit the examination of the entire performance space of the system under examination. Even operating within the same hardware/software environment, it would have been more instructive to examine a greater number of operations using datasets of varying size, density and complexity. However, given the degree of repetition required for each case and the relatively limited testing window available on the FASTPAC service, comparative performance was examined only on a limited number of "representative" operations employing a limited number of data files.

Even with these limitations, the findings should still be of interest to users in a networking environment. Granted, since improvements in speeds across both local- and wide-area networks promise to improve application performance significantly within the near future [Clarkson, 1993], network-related delays may no longer be an issue in the future. However, armed with a more quantitative insight into the types of operations GIS users were invoking and the insights gained concerning FASTPAC performance on these operations, users may gain a better picture of both: (1) how a GIS may perform in such an environment; and (2) which components or processes within the network must be addressed to improve such performance.

In summary, experimental findings indicate that selected GIS operations may be performed across a FASTPAC LAN-interconnect service at speeds comparable to those observed across the individual LANs themselves and that — by extension — the FASTPAC service should be able to deliver satisfactory performance to remote users. However, given both the limitations discussed in previous sections, the growing interest in network-based GIS and the increasing body of knowledge concerning GIS performance testing, there are still important opportunities for further technical research in this field. Selected topics for future research will be introduced and briefly discused in the final section of this chapter.

6.3 SUGGESTIONS FOR FUTURE RESEARCH

The findings from this research point to several problems or phenomena worthy of further investigation. This final section identifies potential areas which merit further research and suggests *if* and/or *how* they may benefit from the research completed here. The following recommendations for further research have been arranged in a very deliberate manner to move from topics arising directly from this research to more general technical subjects which should be addressed in the near future.

(1) The performance monitoring approach used in this research should be extended to examine other factors affecting GIS performance in a client-server environment.

While extensive in their own right, the tests performed only began to define the bounds and characteristics of the performance space of the particular systems in use. This research indicated that the most significant decreases in performance occurred when moving from a stand-alone to LAN environment. Holding the hardware specifications constant, then, it would be instructive to first examine performance tradeoffs in these settings using a wider variety of operations and a greater number of data sets of varying size and complexity.

Other research efforts are now measuring the relative contribution of different hardware components of overall GIS performance (e.g., [Dowers et al., 1990] and [Hawke, 1991]). Using the insights gained from these and the above experiments, it may be worthwhile to then examine the respective impact of hardware enhancements to client and/or server workstations under varying network usage configurations. Such research would help answer important questions concerning the most appropriate and cost-effective approaches to improving GIS performance in client-server environments.

(2) Similar GIS-related performance tests should be conducted across the FASTPAC 2 service to determine the effective differences (if any) which exist between the two services from a user's perspective.

Research indicated that — at least across the Ethernet LANs employed — the full capacity of the 10 Mbit/sec service was rarely utilised for either general UNIX or specific GIS operations. Given the popularity and lower costs of the FASTPAC 2 service, it would be worthwhile to complete similar file transfer, remote login and GIS performance tests across this lower-speed connection to determine how much (if at all) response time degrades in comparison with the FASTPAC 10 results.

(3) Further research is required to determine and minimise the adverse effects of NFS characteristics on inter-network performance.

This research indicated that incremental delays introduced when NFS was called upon to write data onto a remote disk represented a significant component of overall delays across both the LAN and FASTPAC. Research into existing FASTPAC usage trends should determine the extent to which potential clients may use FASTPAC links within an NFS environment. If usage promises to be extensive, the nature of NFS characteristics (particularly the asynchronous NFS *Write* operations) should be further examined as they relate to inter-network performance. Any such research should also consider the potential impact of new hardware extensions (e.g., Sun's *PrestoServe* Board) and planned improvements to the operating systems or network management software in common use.

(4) Subsequent tests should examine other important GIS hardware/software packages commonly found in Australia.

This research only examined the performance and behaviour of one particular GIS package (*Arc/Info*) across an Ethernet LAN. Program structure, memory management, I/O optimisation, and display protocols and processes vary between vendors. For example, as both a hardware and software developer, Intergraph Corporation has developed its own network communication protocols which are supposed to optimise data transmission between Intergraph workstations. Preliminary research within Telecom Australia indicated there was some improvement in GIS performance across FASTPAC using Intergraph's XNS protocols between workstations [Geyman, 1990]. However, those particular tests were designed primarily to confirm the basic viability of the FASTPAC connection on workstations using the XNS protocol and few comparisons between LAN and FASTPAC were even made.

As well, the last two years have seen the emergence of new "review-only" packages which may be further optimised for more efficient data retrieval and display in client-server environments. Simple-to-use map display packages like ESRI's *ArcView* and the *Genabrowse* or *Navigator* packages from GenaSys II Pty. Ltd. are fast becoming standard among end-users to browse GIS database "libraries", select files or datasets of interest and then display them on the screen. If the sales figures for *ArcView* are any indication (10,000 licences sold world-wide in its first six months — [Miller, 1993]), this package alone promises to be used extensively by organisations for the browsing and display of remote files.

With this in mind, it is strongly recommended that similar experiments be performed on other packages to determine if the traffic and performance patterns identified here are commonly encountered or specific to the algorithms and implementation decisions made by ESRI for *Arc/Info.*

(5) The GIS usage monitoring process should be enhanced and extended to: (a) allow more complete reporting of all commands invoked; and (b) identify other statistics (or combinations of statistics) and time-slices which may characterize system usage more appropriately.

Originally begun in 1991, the approach to GIS usage monitoring has already been extended to some degree through 1992 and the flexibility of the summarising process has been significantly enhanced [Morriss, 1993]. Further, the occasional recording problems encountered in pre-Rev. 6.*x* versions of Arc/Info have been apparently addressed by the vendor. Even so, much can still be done in terms of enhancing the basic LOG process within *Arc/Info* to identify and record (e.g.) individual *ArcEdit*, *ArcPlot*, *TIN*, *Network* and *COGO* commands which lay <u>below</u> the top layer of the overall package.

Once this tool has been optimized, a series of systematic, longer-term usage monitoring programs in selected organisations would provide valuable input to determining both application-specific and general changes in GIS usage over time. As well, further development of the framework and heuristics necessary for reliably interpreting these usage statistics is required to better identify requirements for further training, macro command development and potential human-computer interface issues.

(6) An "end-to-end" GIS performance model should be developed to better predict GIS performance and potential communication costs in a client-server environment across both local and wide area networks.

Existing network analysis and design packages now provide estimates of the delays to be expected under varying conditions across networks of varying size and configuration (e.g., [Bachmann et al, 1989]). However, the author could find no commercially-available tools which actually *predicted* the time a particular operation within an application (e.g., a DRAW operation) would take under different conditions. This is not surprising, considering the number of potential combinations of factors which have a significant influence on the response time across a network (i.e., hardware specifications, software, number of users, traffic from other applications on the network, distributed file system characteristics, etc.)

While it would be problematic and time-consuming to develop a complex model which took all potential factors into account, it may be possible to develop a simpler model which takes advantage of the approach employed here to obtain a preliminary series of response-time measurements on specified equipment in both stand-alone and client-server environments. Given this information and employing heuristics which may be gained from other proposed research into the sensitivity of GIS performance with respect to variations in hardware characteristics (e.g., [Wagner, 1991] and [Sloan et al., 1992]), multi-user demands [Dowers, et al., 1990] and data complexity [Hawke, 1991b], a model may be developed which would at least predict how behaviour would change when certain parameters were held constant.

In addition to performance modelling, many of the same modules in such a program could be applied to estimating (and possibly optimising) communications costs within and between sites. Such a process would require input concerning (for example) the locations and mixes of operations present in each different location, the nature and volume of traffic generated by different GIS-related activities, the respective proportions of traffic likely to *stay within* vs. *travel between* LANs and current Telecom rate structures for various services. Using such information, the package would generate estimates of: (a) response-time performance; (b) approximate capital and operating costs involved at each site; and (c) the changes to both as new locations and activities are added and others taken away.

The necessary components and sub-models within the package could be developed and refined using information obtained from: (a) technical research already completed at the University of Tasmania and elsewhere; (b) discussions with Telecom support staff and marketing representatives; (c) Telecom's existing tariff structures; (d) participating organisations; and (e) selected Telecom research projects still to be defined. Rather than developing a package completely from scratch, the prototype package should be built on top of commercially-available network modelling, spreadsheet and/or simulation software.

To summarise, then, the experiments confirmed that selected GIS operations could be performed across a FASTPAC LAN-interconnect service at speeds comparable to those observed across the individual LANs themselves and that — by extension — the FASTPAC service should be able to deliver satisfactory performance to remote users. While further research in specific areas is recommended, these conclusions in themselves may have important ramifications on the operation and management of GIS data and system resources in an organisation.

In fairness, there are important logistical, economic and institutional factors which will have a profound effect on the communications acquisition and implementation strategy of any organisation as well. The final chapter of this dissertation will put this research in a larger context and discuss the potential impact of broadband communications technology on the handling and management of spatial data.

7.0
IMPLICATIONS OF RESEARCH ON GIS OPERATIONS AND MANAGEMENT

This final chapter examines some of the practical considerations affecting the uptake and implementation of broadband communications (and the FASTPAC service in particular) with single organisations. After first outlining the potential impact of broadband communications on GIS-related operations and practices, it then examines some of the factors which may either inhibit or delay the adoption of such services in the near future. Finally, possible refinements and potential topics for future organisational research will be briefly presented and discussed.

7.1 IMPLICATIONS TO THE GIS COMMUNITY

Results of the performance testing indicate that users will have the possibility to access remote disks, processors and output devices and experience little more delay than local users. Ideally, if these files are located in NFS-mounted directories, the differences between accessing remotely-stored files and those stored on a local network should be almost transparent to the end user.

This potentially offers a number of advantages to the GIS user community, including fast display of images and graphics files stored on a remote server, fast transfer of large image and graphics files between servers and diskfull clients, and near-LAN levels of performance when logging-in to remote hosts.

These capabilities, in turn, offer end-users new options not widely found in the GIS community to date. For example:

(1) centralised data management as an alternative to files and databases replicated at each site;

(2) the opportunity to run very complex GIS modelling activities on more powerful computer systems using up-to-the-minute imagery, text and graphics from remote sites (i.e., for real-time modelling of forest fires, flood forecasting, oil-spill monitoring and sea-state conditions).

(3) on-line "browsing" of spatial data and image archives held on remote servers.

Some of these options are being introduced and tested on a limited basis already using satellite-based communications or lower-speed, cable-based services. For example, the potential opportunities and constraints involved in providing on-line systems for the browsing and retrieval of satellite imagery are now being investigated, with prototypes being developed in both the United States [Dozier, 1991] and Canada [Whitehouse, 1993].

Despite the visions of multi-participant, jurisdiction-wide efforts which have dominated much of the rhetoric in the spatial data handling community (See Section 2.2.1), much of the initial interest in broadband services will probably come from individual organisations. Recognised hierarchies and clearer lines of authority in a single organisation mean that strategic decisions regarding the acquisition of higher-speed networking services may potentially be taken over a much shorter period.

Once inside, the manner in which these individual organisations involved with GIS manage their information will have also a bearing on their communication requirements. Table 7.1 compares the potential for FASTPAC service within the framework developed earlier and suggests factors which should be considered in each category.

GIS software packages may be found in many different organisations supporting many different applications. However , *major GIS databases* and *digital map* *libraries* are still predominantly found supporting long-term programmes in large, centralised organisations with clearly-defined mandates (e.g., utilities, municipalities, forest management organisations and government mapping agencies). These organisations have generally made a substantial investment in system management and database maintenance, and must eventually consider the costs of establishing database or map library access from remote locations.

By comparison, many other organisations tend to use GIS in support of shorter-term projects with a focused set of objectives and a limited life-span. Research organisations and educational institutions clearly fall into this category, as well as some engineering design, environmental assessment and land-use planning units within government. In such cases, there is less need, opportunity or even desire for centralised processing or data storage. Even in cases where there is a clear requirement for occasional access to outside data sources, many of these organisations (especially in research and education) use AARNet links for low-cost data transfer.

In such cases, it is likely that GIS requirements would only represent a minor component of the overall justification of FASTPAC in an organisation. Unless it could be confirmed that the users in question were willing to "lead the charge" and exercised a great deal of influence in the organisation, it is likely that any procurement decision will be based on the wider systems management and communications requirements of the organisation.

Table 7.1
FASTPAC Applicability to Different Data Management Strategies

Approach	Degree of Centralisation in:		FASTPAC Applicability
	Processing	Storage	
1	Decentralised	Decentralised	FASTPAC well-suited to this environment. • "Browsing" of remote files prior to retrieval; • High-speed file transfer between sites. *However,* watch for situations where: • staff in every location already possess the specific data required to carry out their own tasks (hence little routine data transfer); and/or • Graphics and attribute data files do not change frequently.
2	Decentralised	Centralised *(All data stored on central file server)*	Ideal environment for FASTPAC, since corporate information management policy dictates: • Routine copying of selected files between central site and local workspace; • Possible processing, edit, analysis and display of NFS-mounted files. *N.B. Above caveats apply here as well.*
3	Centralised	Decentralised *(Local storage possible; only corporate data stored on central server)*	Unusual situation. Probably not suitable for FASTPAC except where GIS application software is stored centrally or where remote users access high- performance computers. Applicability of FASTPAC in each location will depend on: • the extent to which corporate data is routinely accessed by local users; • the relative degree of "perishability" of the corporate data (i.e., frquency of changes to corporate data files).
4	Centralised	Centralised	FASTPAC would deliver excellent response times, but its cost-effectiveness in (particularly) environments of continuous connection may be questionable. Intermittent bursts in network traffic would be due to displays of graphics and image files. Frequency of such displays unclear, and may reduce as workstations and X-terminals come with more memory. *Additional research required to determine FASTPAC cost-effectiveness in Client/Server environment using X-Terminals.*

Longer-Term Effects

While identifying near-term applications (like those described above) of broadband communications technology on GIS operations and management may be relatively straightforward, the longer-term effects will be more difficult to predict. For example, the photocopier, the fax machine and even television programming are all examples of products or services originally designed with one market or application in mind that now being applied to many different situations. The principal use of the Internet itself, while originally designed for on-line connection to supercomputer installations, has evolved substantially over the past twenty years; the current interest in the Internet by government and commercial users around the world has much more to do with widespread electronic mail, bulletin boards and file transfer capabilities than with supercomputer access [Sproull et al., 1991].

Recent reports have indicated that information technology is finally demonstrating a net positive effect on productivity in the workplace. However, it is not because computers and networks have made existing work more efficient; rather, products and processes are being completely redefined or "re-engineered" to take better advantage of the processing and communications capabilities of today's technology *[Business Week*, 1993].

The various means in which broadband communications technology may eventually contribute to the re-engineering of GIS usage in an organisation remain unclear. On the software side, public broadband networks may ultimately enable on-line diagnosis of software problems at the user's site by remote vendors and fast downloading of subsequent software revisions to end-users. On another front, the Canadian Hydrographic Service is now beginning to examine the communications capabilities and operational strategies required for faster turnaround times for updates to electronic charts using high-speed networks to link production, quality control and distribution units together [Anderson, 1993]. Given the growing popularity of the Lotus *Notes* software for data distribution and collaborative document preparation within a workgroup [Marshak, 1992], there may eventually be collaborative versions of major GIS and mapping packages as well.

Before these prospects reach the implementation stage, however, there are several practical and/or economic implications which must be overcome. Key factors which may inhibit the fast adoption of broadband communications networks will be discussed in the next section.

7.2 CONSTRAINTS TO NETWORK IMPLEMENTATION

Previous sections have provided an idea of both the performance and the potential applications of broadband communications to spatial data handling and management. To be realistic, however, many technical, operational and institutional problems involved in providing consistent and ready access to multiple databases – and then ensuring standard and well-documented "views" of the data and resources – continue to impede the progress of many organisations toward such goals. A full discussion of the issues and problems related to telecommunications may be found in [Newton et al., 1992a]. Some of the key issues include:

(1) Availability of Communications Infrastructure

(2) Potential Cost of Service

(3) Security Concerns

(4) Potential Inability to Sustain Long-Term Management Support

(5) Lack of Progress in Adopting Common Data Standards

(6) Lack of User Awareness

Availability of Communications Infrastructure

In addition to the fast packet switching hardware necessary for FASTPAC use, such high-speed telecommunication also requires access to a *fibre optic infrastructure*. This entirely new infrastructure – just now being put in place in major Australian centres – is central to FASTPAC's 'high speed' service offering. In most developed countries, the fibre optic cabling necessary to support high-speed communications services is still not in place outside metropolitan areas or inter-city corridors.

The speed with which the fibre optic MANs are established in each metropolitan area, together with their location relative to potential customers, will affect the widespread adoption of advanced communication services. Telecom Australia is now carrying out an optical fibre cable replacement programme, and, over the next few years, several million homes and businesses throughout Australia will find themselves within 2 kilometres of optical fibre service [Hunter, 1991]. All the same, the incremental cost and time involved to bring fibre into all homes and offices may be significant. Creating the market necessary to justify such a move will depend on the ability of utilities to offer new, video-based services to their customers, including video-on-demand, in-house shopping, cable-based meter reading and even in-house video gambling..

Potential Cost of Service

The issue of whether organisations will be willing to pay extra for broadband services will depend on how quickly users can productively apply the full range of benefits such technology offers. ISDN market studies from Telecom France in 1988 indicated that – after a year of commercial operation – batch file transfers and routine back-ups still constituted over 80% of the tasks assigned to ISDN networks [Zwart et al, 1992]. Further, experiences and market surveys indicated that no more than 10 per cent of Telecom France customers already using leased lines or public packet networks would be prepared to switch to ISDN at that time if related operating expenses increased by more than 50%.

A review of the Telecom FASTPAC 10 rate structures indicates that customers wishing to link together two sites across a metropolitan area would invest $18,000 in installation fees and approximately $116,000 in fixed annual access and interface rentals before transferring any data. Even beginning at minimal levels of data transfer per month, they would likely budget another $5100-$10,200 per year for volume link charges for local connections, and over $60,000 per year for longer-distance links. (Note: Details concerning the cost of operations within the FASTPAC rate structure as it appeared in September, 1992 may be found in Appendix E.)

While these expenses are significant, they are roughly equal to the fully-loaded costs of 3-4 extra systems support persons. If an organisation is considering replicating and placing a copy of all or some of its various databases in remote offices to improve performance, then the costs of the additional support required for such an activity should be compared to the costs of a FASTPAC link.

Concerns over System and Data Security

Informal exchange of non-confidential hardcopy information within and between organisations has been common for years at operational levels. Gaining on-line access to an organisation's *computer files*, however, introduces new problems which often demand more complex technical solutions and require more formal arrangements between the parties involved [Anderson, 1992]. Further, customers may be understandably reluctant to send confidential corporate data over public networks unless some degree of data encryption is in place.

Concerns over Longer-Term Management Support

Even after information-sharing agreements and policies are in place, programme managers face a serious challenge in winning and keeping the long-term management support required to fund physical network development and on-going operations. Proponents of state-wide networks, for example, are often hard-pressed to present good business cases for development of such networks. For example, IDON [1990] cited one situation where: (a) the lack of management awareness and experience with networking opportunities; (b) the high capital costs involved; and (c) the lack of any single predominant application justifying network implementation by itself discouraged managers from supporting a jurisdiction-wide networking efforts at that time.

Efforts within a single organisation are often easier to justify. Tangible costs and benefits can usually be more readily identified, and organisations may often link their data- and voice-communication requirements together when planning their future communication strategies [Chan, 1991].

Lack of Progress in Adopting Common Data Standards

The incompatibilities existing between hardware, software and databases greatly increases the costs involved in putting a "spatial information infrastructure" or "virtual database" in place. Craig et al. [1991], for example, documents specific problems involved in transferring data between systems using direct translation and established interchange formats. Despite the early recognition of this problem, it continues to worsen as the number of vendors increases, existing systems evolve, user communities begin to diverge and a limited number of applications gain wider, "mass market" appeal.

The standards required for successful implementation of spatial information networks in fact go well beyond the problems of file transfer between systems. Coleman and McLaughlin [1991], for example, suggest the standards now affecting the spatial data community can be grouped into four categories, including Hardware & Communications, Software, Data Specifications & Formats, and Standard Data Sets. While it is too small at present to significantly influence decisions in the first two categories mentioned, the spatial information community *can* exercise significant control over the remaining ones.

The problem in the community today is not a lack of such standards; rather, too many options are being developed and proposed as "national" or "international" standards. For example, Clarke [1992] describes one such standard -- the Spatial Data Transfer Standard (or SDTS) -- and the long efforts to have it approved as a national standard in the United States. Parallel development and testing efforts are also underway in Canada [Sondheim, 1991] and elsewhere, but these propose alternative products and approaches which treat such aspects as data modelling, encoding and documentation in a different manner.

In the absence of any single predominant standard, some organisations and even jurisdictions have tried to minimise incompatibilities between working groups by specifying "corporate guidelines" for system acquisition, operating procedures and even definition of key types of data. Recent U.S. Federal Government purchases suggest a more single-minded approach to the purchase of GIS-related hardware and software, and many jurisdictions are now taking a more proactive role in reaching consensus on standards [Coleman and Ogilvie, 1991].

While progress at these levels is important and encouraging, the lack of agreement on international standards will still slow the widespread acceptance of value-added application software, network access tools, and information products in electronic form.

Lack of User Awareness

Even three years after the Telecom marketing study was undertaken, one of the main inhibitors to the uptake of broadband services remains the lack of awareness among users. Many potential users — especially those still unfamiliar with wide area electronic mail, file transfer and other capabilities offered through the Internet (and equivalent services) — simply are still unaware of the sweeping changes to the way organisations are now conducting business.

All six of these constraints represent significant impediments to the widespread acceptance and adoption of broadband networks in the spatial data handling community, and their potential influence should not be underestimated. In fairness, however, market demands are evolving rapidly as the technology improves, competition increases and prices drop. For example:

(1) the 80% per year annual growth rate in LAN users indicates a solid corporate acceptance of network computing for both local operations;

(2) the growing demand to interconnect local area networks within single organisations has already been recognised as a major market for FASTPAC [Montgomery, 1992];

(3) the easing of access restrictions and more effective promotion of Internet services has already lead to a phenomenal 20% *per month* growth in the number of Internet users over the past 24 months [Baker, 1993].

(4) the upper bounds of network performance and capacity are finally reaching the point where they can effectively accommodate remote access to larger graphics and image datasets in real time; and

(5) the increasing use of electronic mail and the emergence of collaborative, "workgroup computing" software are both fundamentally changing the manner in which people communicate and work with one another.

Among both suppliers and corporate users, levels of expectation with respect to networking performance and functionality have dramatically increased over the past two years. As with computer processing power, "fast" will never be "fast enough" when it comes to network performance. As (a) the market becomes more accustomed to the range of services now being offered or proposed; (b) users begin to depend more and more on larger and more complex data files to support routine operations; and (c) suppliers gain more feedback in how to best package their products and services, the demand for broadband communications services within and between organisations will certainly increase.

Whether this increased demand and functionality will move the majority of the user community back towards a centralised model for data storage remains to be seen. While the technology itself could support such a move — and while it may represent a more efficient use of both equipment and technical staff — such a decision rests more with individuals than with equipment. As corporate and national GIS data holdings eventually blend together across broadband networks into a global spatial data infrastructure, the philosophies affecting the location and management of individual databases may become a moot point so long as the user can gain quick and transparent access to the information and services desired.

7.3 CONCLUDING REMARKS

In the period since this research has been carried out, an increasing number of articles in technical journals, business magazines and even the popular press have discussed the promise of broadband communications and its potential application to "video-on-demand", personal communications, on-line libraries and even to "video-on-demand", personal communications, on-line libraries and even "electronic town hall meetings". The *Sequoia 2000* project has begun to examine the potential application of GIS and broadband communications technology to global change monitoring [Stonebreaker, 1992]. There has even been renewed enthusiasm for the planning and establishment of a "National Spatial Data Infrastructure" in the United States [Mapping Sciences Committee, 1993].

These articles and initiatives have played an important role in developing a heightened awareness of networking possibilities in different business and professional communities. In light of this growing appreciation of current networking capabilities and the growing number of applications designed to take advantage of networked resources, a "grass-roots" demand for services is quickly growing to match to "top-down" networking visions being touted through the 1980's. As Sproull et al. [1991] pointed out, it may take years to identify and fully appreciate the "second-order" or indirect effects of such high-speed communications technology even within the spatial data handling community, much less the mass market.

It is unlikely that an organisation's GIS or remote sensing requirements alone would justify the acquisition and implementation of broadband communications yet — few executives would find it cost-effective to apply an expensive service to the rarefied and relatively low-profile functions of a spatial data handling unit alone. However, it *is* likely that large organisations in Australia, North America and Europe will acquire such technology to fulfil most of their internal communication requirements *and* that GIS users will be able to take advantage of this technology within their organisation in the very near future. Given the results of these experiments, spatial data managers should promote and support the careful planning and implementation of such technology in their organisations to ensure that their unique requirements are properly taken into account.

Bibliography

Alexander, D. and M. Fox [1992]. "Statewide Networking of Natural Resources Information." in *Networking Spatial Information Systems*, P. Newton, P.Zwart and M.Cavill, eds. Belhaven Press, London.

ALIC [1990]. *National Strategy on Land Information Management.* Australian Land Information Council.

Amundsen, S. [1989]. "Performance Testing of a Gridcell-Based GIS". *Proceedings of thje 9th International Symposium on Computer-ASssisted, Cartography (AUTO-CARTO 9),* American Society for Photogrammetry and Remote Sensing, Falls Church, Virginia, U.S.A.

Andersen, N.M. [1990]. "ICOIN Infrastructure". *Proceedings of a Forum on the Inland Waters, Coastal and Ocean Information Network (ICOIN)*, pp. 4-17. The Champlain Institute, Fredericton, New Brunswick, Canada. June.

Anderson, H. [1992]. "Networking: A View from the Next Century." *Data Communications*, pp. 129-134, September.

Anderson, R. [1992]. "Access and Privacy of Distributed Land Related Information" Unpublished MScE. thesis, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B., Canada.

Antenucci, J. and K. Brown and P.L. Crosswell and M.J. Kevany [1991]. *Geographic Information Systems: A Guide to the Technology.* Van Nostrand Reinhold, New York, New York, U.S.A.

Archer, H. and A. Ahner [1990]. "Bibliography on Economic Cost/Benefit Aspects of GIS and AM/FM Projects." Internal document available through Plangraphics Inc., Frankfort, KY, U.S.A.

Aronson, P.B. [1982]. "A Comparative Analysis of Three Polygon Overlay Algorithms." Unpublished M.Sc. Thesis, Department of Geography, State University of New York at Buffalo, Buffalo, N.Y., U.S.A.

Bachmann, D.W. and M.E. Segal and T.J. Teorey [1989]. "On the Performance of Copying Large Files Across a Contention-Based Network". Technical Report CITI-TR-89-3, Center for Information Technology Integration, University of Michigan, Ann Arbor, Michigan, U.S.A. June.

Bachmann, D.W. and M.E. Segal and and M.M. Srinivansan and T.J. Teorey [1989]. "NetMod: A Design Tool for Large-Scale Heterogeneous Campus Networks". Technical Report CITI-TR-90-1, Center for Information Technology Integration, University of Michigan, Ann Arbor, Michigan, U.S.A. June.

Bagnara, S. and A. Rizzo [1989]. "A methodology for the analysis of error processes in human-computer interaction". In: *Advances in human factors/ergonomics 12A -Work with computers: Organisational, management, stress and health aspects.* Vol. 1 of Proceedings of the Third International Conference on Human-Computer Interaction, Boston. Smith, M.J. and G. Salvendy (eds) Elsevier Science, Amsterdam. pp 605-612.

Beker, D. [1993]. "The Environmental Information Network: From Concept to Strategy". Contract Report prepared for the State of the Environment Reporting Branch, Environment Canada, Ottawa, Ontario, Canada. March.

Bell, K. [1988]. Land Information Integration: The Hub System in Queensland (March 1988)." *Proceedings of the 1988 Annual Meeting of the Urban and Regional Information Systems Association (URISA].*, Vol. III, pp. 165-180. Urban and Regional Information Systems Association, Washington, D.C., U.S.A.

Bennett, P.G.L. [1991]. "Distributed Processing Pilot: Integrating Spatial and Textual Data Residing on Multiple Platforms." *Proceedings of the 1991 Annual Conference of the Urban and Regional Information Systems Association* (URISA), vol. 3, pp. 25-38.

Berg, K. and W. Boebert and W. Franta andf T. Moher [1982]. *Formal Methods of Program Verification and Testing.* Prentice-Hall, Englewood Cliffs, N.J., U.S.A.

Boggs, D.R. and J.C. Mogul and C.A. Kent [1988]. "Measured Capacity of an Ethernet: Myths and Reality." WRL Research Report 88/4. DEC Western Research Laboratory, Palo Alto, CA., U.S.A. September.

Branscomb, A. [1982] "Beyond Deregulation: Designing the Information Infrastructure." *The Information Society Journal*, Vol. 1, No. 3, pp. 167-190.

Brown, K. [1991]. "Political Quagmire Stymies Technology's Promise." Geo-Info Systems, Vol. 1, No. 10. November/December.

Business Week [1993]. *Business Week* Magazine: Special Issue on "The Technology Payoff", pp. 56 - 79. June 14, 1993.

Cabrera, L-F and D.D.E. Long [1991]. "Exploiting Multiple Bit Streams to Provide High Data Rates." *Proceedings of the 1991 Summer Usenix Conference*, Nashville, Tenessee, Usenix Association, June.

Cady, J. and B. Howarth [1990]. *Computer Systems, Performance Management and Capacity Planning.* Advances in Computer Science Series. Prentice-Hall of Australia, Sydney, N.S.W., Australia.

Calkins, H. [1983]. "A Pragmatic Approach to Geographic Information Systems Design". in.*Design and Implementation of Computer-Based Geographic Information Systems.* IGU Commission on Geographical Data Sensing and Processing: Amherst, New York, U.S.A.

Camarata, S. [1992]. "The Integration of New Technologies and Distributed Architecture: A New Revolution for Geographic Information Systems". in *Networking Spatial Information Systems*, P.W. Newton, P.R. Zwart and M.E. Cavill, eds. Belhaven Press, London.

Cavill, M. and M. Montgomery [1991]. "The Role of Telecommunications in GIS in the 1990's and Beyond." Paper presented at the Longman Professional Conference on Geographical Information Systems, Melbourne, Victoria, June.

Cavill, M. and P. Kirton and P.R. Zwart [1989]. "Networking GIS Users – The Impact of Broadband Networks." *Proceedings of URPIS 17.* Australian Urban and Regional Information Systems Association, Perth, pp. 406-416.

Champine, G.A., R.D. Coop and R.C. Heinselman [1980]. *Distributed Computer Systems.* North Holland Publishing Company, Amsterdam.

Chan, C.W.Y. [1991]. "Performance Analysis of GIS Based on Different Data Characteristics." Unpublished M.Sc. Thesis, Department of Geography, University of Edinburgh, Edinburgh, Scotland, U.K.

Chan, L. [1991]. Personal communication with Manager of Computing and Communications for Victoria Department of Conservation and Environment, Melbourne, Victoria, August.

Chapman, G.M. [1991]. "Integrating Facilities with a Land Information System." *Proceedings of the 1991 A/E/C Conference*, Sydney, N.S.W., pp. 11-15.

Chorafas, D.N. [1980]. *Computer Networks for Distributed Information Systems.* Petrocelli Books, New York, N.Y., U.S.A..

Christie, D. [1984]. The Development of Theme Overlay Capability for CARIS." Unpublished MEng. Report, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B., Canada.

Clarke, Drew [1990]. "A Standard Format for Spatial Data Transfer." *Proceedings of the Workshop on Networking Spatial Information Systems.* Telecom Australia Research Laboratories, Geographic Information Studies Group, Clayton, Australia. May.

Clarkson, M.A. [1993]. "Hitting Warp Speed for LANs". *Byte* Magazine, Vol. 18, No. 3, pp. 123-128. March.

Coleman, D.J. [1988]. "Implementing a Land Information Network in New Brunswick". Unpublished MScE. thesis, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B., Canada.

Coleman, D.J. [1991a]. Internal project memo discussing proposed procedures for characterising GIS usage. Centre for Spatial Information Studies, University of Tasmania, Hobart, Tasmania, Australia.

Coleman, D.J. [1991b]. GIS/FASTPAC Network Monitoring Experiment: Proposed Design." Internal project report. Centre for Spatial Information Studies, University of Tasmania, Hobart, Tasmania, Australia.

Coleman, D.J. [1992a]. "GIS Performance in a Broadband Communications Environment." *Contract report to Fast Packet Services Group, Telecom Australia, Melbourne.* Centre for Spatial Information Studies, University of Tasmania, Hobart, Tasmania. February.

Coleman, D.J. [1992b]. "Final Report: 1992 GIS/FASTPAC Performance Testing and Network Monitoring Experiments." *Contract report to Fast Packet Services Group, Telecom Australia, Melbourne.* Centre for Spatial Information Studies, University of Tasmania, Hobart, Tasmania. October.

Coleman, D.J. and J.D. McLaughlin [1991a]. "Standards for Spatial Information Interchange: a Management Perspective." *The CISM Journal.* In Print. Canadian Institute of Surveying and Mapping, Ottawa, Canada.

Coleman, D.J. and M. Ogilvie [1991b]. "Reaching Consensus on Spatial Data Standards at the State/Provincial Level: A Canadian Example." *Proceedings of the 1991 Australasian Urban and Rural Information Systems Association* (AURISA), Wellington, New Zealand. November.

Coleman, D.J. and P. Morriss and P.R. Zwart [1992a]. "Quantifying GIS Usage in an Organisation: Developing Management Tools". *Proceedings of the 1992 Australasian Urban and Regional Information Systems Association (AURISA) Conference,* pp. 27-36. Brisbane, Queensland, Australia. November.

Coleman, D.J. and P.R. Zwart [1992b]. "Tradeoffs Towards Real-time Spatial Information Networks." in *Networking Spatial Information Systems,* P.W. Newton, P.R. Zwart and M.E. Cavill, eds. Belhaven Press, London.

Coleman, D.J. and P.R. Zwart [1992c]. "Modelling Usage and Telecommunications Performance in Real-Time Spatial Information Networks". *Proceedings of the 5th International Symposium on Spatial Data Handling,* Vol. 1, pp. 144-153. Charleston, South Carolina.

Coleman, D.J. and P.R. Zwart [1992d]. "Determining GIS Performance in a Broadband Telecommunications Environment". *Proceedings of the 1992 Australasian Urban and Regional Information Systems Association (AURISA) Conference,* pp. 421-431. Brisbane, Queensland, Australia. November.

Comer, D. [1988]. *Internetworking with TCP/IP: Principles, Protocols and Architecture.* Prentice-Hall, Englewood Cliffs, New Jersey, U.S.A.

Cook, G. [1989]. "An Enterprise-Wide GIS: The Integration of GIS with I/S." *Proceedings of the 1989 Annual Meeting of the Urban and Regional Information Systems Association (URISA].,* Vol. III, pp. 42-52. Urban and Regional Information Systems Association, Washington, D.C., U.S.A.

Cook, R. and J. Kennedy [1967], editors. *Proceedings of a Tri-State Conference on a Comprehensive, Unified Land Data (CULDATA) System, December* 9-10, 1966. University of Cincinatti, Cincinatti, Ohio, U.S.A.

Craig, W.J. and P. Tessar and N. Ali Khan [1991]. "Sharing Graphic Data Files in an Open System Environment." *URISA Journal,* vol. 3, no. 1, pp. 20-32. Spring.

Crosswell, P.L. [1986]. "Developments in Data Transfer Between Geographic Information Systems and Mainframe Computer Databases." *Proceedings of the 1986 Annual Conference of the Urban and Regional Information Systems Association.*

DeWitt, D.J. [1985]. "Benchmarking Database Systems: Past Efforts and Future Directions." *Database Engineering*, Vol. 8, No. 1, pp. 2-9.

Diebold, J. [1979]. "Information Resource Management: The New Challenge." *Infosystems*, June, pp.50-53.

Dowers, S. and B.M. Gittings and T.M. Sloan and T.C. Waugh and R.G. Healey [1990]. "Analysis of GIS Performance on Parallel Architectures and Workstation-Server Systems." *Proceedings of the 1990 LIS/GIS Symposium*, pp. 555-561.

Dozier, J. [1991]. "Early EOSDIS: A Data and Information System for the Study of Global Change," Sequoia 2000 Technical Report 91/8, University of California, Berkeley, CA, U.S.A. December.

Edwards, D.R. and T.W. Foresman [1990]. "Centralized Management: Protecting Your Data Investment." *Proceedings of the 1990 Annual Meeting of the Urban and Regional Information Systems Association (URISA).*, Vol. II, pp. 32-44. Urban and Regional Information Systems Association, Washington, D.C., U.S.A.

Egan, B.L. [1991]. *Information Superhighways: The Economics of Advanced Public Communications Networks.* Artech House, Norwood, Massachusetts, U.S.A.

ERIN [1991]. "ERIN: A Communications Network." Internal Implementation Planning Document prepared for the Environmental Resources Information Network (ERIN). Ministry of the Arts, Sports, the Environment, Tourism and Territories, Commonwealth Government of Australia, Canberra.

Ezigbalike, I.F.C. [1988]. "Land Information System Development: Software and Management Considerations." Ph.D. Thesis, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B., Canada. June.

Fenwick, P. [1990]. "Performance Measurements on a 39 km. Extended Ethernet." *The Computer Journal.*

Ferrari, D. [1978]. *Computer Systems Performance Evaluation.* Prentice Hall, Englewood Cliffs, New Jersey, U.S.A.

Ferrari, D. and G. Serazzi and A. Zeigner [1983]. *Measurement and Tuning of Computer Systems.* Prentice-Hall, Englewood Cliffs, New Jersey, U.S.A.

Ferrari, D. and J. Pasquale and G.C. Polyzos [1991]. "Network Issues for Sequoia 2000." Sequoia 2000 Technical Report 91/6, Computer Science Division, Department of Electrical Engineering and Computer Science, University of California, Berkeley, California, U.S.A.

Finnie, G. [1992]. "The Rise and Fall of B-ISDN". *Communications Week International*, June 22, 1992, pp. 18-19.

Flury-Herard, B. and J-P. Temime [1988]. "Evaluation of the ISDN Market in France." Proceedings of the IEEE Conference on ISDN Telecommunications. IEEE Publication No. CH2536-1/88/0000-0066.

Forrest, D. [1992]. "Implementing a Provincial Land Information Network." in *Networking Spatial Information Systems*, P. Newton, P.Zwart and M.Cavill, eds. Belhaven Press, London.

Fowler, H.J. and W.E. Leland [1991]. "Local Area Network Traffic Characteristics, with Implications for Broadband Network Congestion Management." *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 7, pp. 1139-1149. September.

Fox, M. [1992]. Personal communication with Spatial Systems Group Manager, Victoria Department of Conservation and Environment, Melbourne, Victoria. June.

Geist, R. and R. Allen and R. Nowaczyk [1987]. "Towards a Model of User Perception of Computer Response Time", *CHI'87 Proceedings*, pp. 249-254. Association of Computing Machinery (ACM). May.

Goldmark, P.C. [1972]. "Communication and the Community." *Scientific American*. Special Issue on Communications. September.

Goodchild, M.F. and B.R. Rizzo [1986]. "Peformance evaluation and workload estimation for geographic information systems." *Proceedings of the 2nd International Symposium on Spatial Data Handling,* pp. 497-509. Seattle, Washington, U.S.A., August.

Greenwell, J. [1992]. "Communications Issues for Land Information Sytstems". *Seminar Notes and Materials, 1992 Workshop on Networking Spatial Information Systems.* Centre for Spatial Information Studies, University of Tasmania, Hobart, Tasmania, Australia.

Guevara, J.A. [1983]]. "A Framework for the Analysis of Geographic Information Systems Procedures: The Polygon Overlay Problem, Computational Complexity and Polyline Intersection." Unpublished Ph.D. Dissertation, Department of Geography, State University of New York at Buffalo, Buffalo, N.Y., U.S.A.

Geyman, V. [1990] Internal memo on FASTPAC/GIS demonstration results. Fast Packet Services Group, Telecom Australia, Melbourne, Vic.

Gunton, T. [1989]. *Infrastructure: Building a Framework for Corporate Information Handling.* Prentice-Hall International (U.K.) Ltd., Hertfordshire, U.K.

Hamilton, A.C. and A. Chrzanowski and N. McNaughton [1976]. "Infrastructure Information Requiremnents in the Maritime Provinces: An Analysis" Technical Report No. 37, Department of Surveying Engineering, University of New Brunswick, Fredericton, New Brunswick, Canada. June.

Hammer, S.W. and R.L. Clay [1992]. "Would I Buy Servers All Over Again? Performance Matters in a UNIX Environment". *Proceedings of the Twelfth Annual ESRI Users Conference*, pp. 241-252. Environmental Systems Research Institute, Redlandfs, CA., U.S.A.

Hawke, D. [1991]. "Characterising the Performance of Geographic Information Systems." *Proceedings of the Third Annual Colloquium of the Spatial Information Research Centre*, University of Otago, Dunedin, New Zealand. May.

Hawke, D. [1991]. "The Influence of Data Complexity on GIS Performance". *Proceedings of the 19th Australiasian Conference on Urban and Regional Information Systems*, pp. 57-65. Wellington, New Zealand. November.

Hawthorn, P. [1985]. "Variations on a Benchmark." *Database Engineering*, Vol. 8, No. 1, pp. 19-28.

Healey, R.G. [1993]. Personal communication, Department of Geography, University of Edinburgh, Edinburgh, United Kingdom. September.

Healey, R.G. and S. Dowers, B.M. Gittings, T.M. Sloan and T.C. Waugh [1991]. "Determination of Computing Resource Requirements for GIS Processing in a Workstation-Server Environment." *Proceedings of EGIS '91 Conference*, Brussels, Belgium, 2nd April, 1991

Hearle, E.F.R. [1962]. "A Data Processing System for State and Local Governments." *Public Administration Review*, Vol. 22, No. 3, pp. 146-152. September.

Heidelberger, P. and S. Lavenberg [1984]. "Computer Perormance Evaluation Methodology". *IEEE Transactions on Computers*, Vol. C-33, No. 12, pp. 1195-1220.

Hewlett-Packard [1989]. *Users Guide: HP18212A — LAN Performance Analysis Application Software.* User Manual Part No. 04972-99902. Telecommunicatons Division, Hewlett-Packard Co. Ltd., Colorado Springs, Colorado, USA.

Howes, D.A. [1991]. "Performance Characteristics of Triangulation Algorithms." Unpublished M.Sc. Thesis, Department of Geography, University of Edinburgh, Edinburgh, Scotland, U.K.

Hui, J.Y. [1990]. *Switching and Traffic Theory for Integrated Broadband Networks.* Kluwer Academic Publishers, Norwell, Massachusetts, U.S.A.

Hunter, G. [1991]. "Potential of Spatial Information Science." Search, Vol. 22, No. 2. March.

IDON Corporation [1990]. "Final Report: A Federal Geographic Information Systems (GIS] Network Requirements Study." Contract Report prepared for the Data Communications and Networking Subcommittee, Interagency Committee on Geomatics, Government of Canada. Ottawa, Canada. November.

Ingoldsby, T.R. [1991]. "Transparent Access to Geographically Related Data from Heterogeneous Networked Systems." *Proceedings of the 1991 Annual Conference of the Urban and Regional Information Systems Association* (URISA), vol. 3, pp. 14-24.

Jain, R. [1991] *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling.* Wiley. New York, New York, U.S.A.

Katz, R. and G. Gibson and D. Patterson [1989]. "Disk System Architectures for High Performance Computing." *Proceedings of the IEEE, Special Issue on Supercomputing*, December.

Katz, R. [1991]. "High Performance Network and Channel-Based Storage." Sequoia 2000 Technical Report 91/2, Computer Science Division, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA., U.S.A.

Keen, P.G.W. [1991]. *Shaping the Future: Business Design Through Information Technology.* Harvard Business School Press, Boston, MA., U.S.A.

Kirton, P. [1990]. "ISDN and B-ISDN Networks." *Proceedings of the Workshop on Networking Spatial Information Systems.* Telecom Australia Research Laboratories, Geographic Information Studies Group, Clayton, Australia. May.

Lam, N. [1977]. "Polygon Overlay: An Examination of an Algorithm and Related Problems". Unpublished Master's Thesis, Department of Geography, University of Western Ontario, London, Ontario, Canada.

Langille, D. [1991]. "Progress Report on LRIS Gateway Development." *LRIS Newsletter*, vol. 12, no. 5, pg. 2. September. Land-Related Information Services Branch, Alberta Department of Forestry, Lands and Wildlife, Edmonton, Alberta, Canada.

Lau, S.L. [1991]. "Disk, Network and Operating System Effects on GIS Performance." Unpublished M.Sc. Thesis, Department of Geography, University of Edinburgh, Edinburgh, Scotland, U.K.

Loukes, D.K. and N. Nandlall [1990]. "Developing a Land Information Network for New Brunswick: Implementation Recommendations." *Proceedings of the 1990 Canadian National Conference on Geographic Information Systems*, pp. 17-40. Canadian Institute of Surveying and Mapping, Ottawa, Canada.

Mapping Science Committee [1990]. *Spatial Data Needs: The Future of the National Mapping Program.* Report of the Mapping Sciences Committee, Board on Earth Sciences and Resources, Commission on Physical Sciences, Mathematics and Resources, National Research Council. National Academy Press, Washington, D.C., U.S.A.

Mapping Science Committee [1993]. *Towards a Coordinated Spatial Data Infrastructure for the Nation.* Report of the Mapping Sciences Committee, Board on Earth Sciences and Resources, Commission on Geosciences, Environment and Resources, National Research Council. National Academy Press, Washington, D.C., U.S.A.

Marble, D.F. and E.A. Wentz [1989]. "The Development of Standardized Benchmarks for Geographic Information Systems: A Progress Report." *Proceedings of the 1989 ESRI User Conference,* Palm Springs, CA.

Marble, D.F. and L. Sen [1986]. "The Development of Standardized Benchmarks for Spatial Database Systems". *Proceedings of the 2nd International Symposium on Spatial Data Handling,* pp. 488-496. Seattle, Washington, U.S.A., August.

Marshak, D. [1992] "The Notes Phenomenon: The Industry Reacts to Lotus Notes". *Office Computing Report*, Vol. 15, No. 12, pp. 3 - 24. Patricia Seybold Group, Boston, Massachusetts, U.S.A. December.

Mayhew, D.J. [1992]. *Principles and Guidelines in Software User Interface Design.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, USA.

McHarg, I.L. [1969]. *Design with Nature.* Doubleday Books, New York, New York, U.S.A.

McLaughlin, J.D. [1991a]. "Towards National Spatial Data Infrastructure." *Proceedings of the 1991 Canadian Conference on GIS*, Ottawa, Canada, pp. 1-5. Canadian Institute of Geomatics, Ottawa, Canada. March.

McLaughlin, J.D. [1991b]. "National Spatial Data Infrastructures: The Next LIM Challenge." *Proceedings of the 1991 NSW Land Information Management Conference*, pp. 159-166. Monograph No. 14, School of Surveying, University of New South Wales, Kensington, NSW, Australia.

McNair, E.A. and G.H. Sauer [1985]. *Elements of Practical Performance Modelling.* Prentice Hall, Englewood Cliffs, N.J., U.S.A.

Mendenhall, W. and L. Ott and R. Scheaffer [1971]. *Elementary Survey Sampling.* Duxbury Press, Wadsworth Publishing Company, Inc., Belmont, Aalifornia, U.S.A.

Miller, A. [1990]. "GIS in a Networking Environment." *Proceedings of GIS'90, International Symposium on Geographic Information Systems.* Forestry Canada. pp. 255-262.

Miller, A. [1993]. Personal communication with the President of ESRI Canada. March.

Montgomery, M. [1992]. Personal communication with Manager of Marketing, Fast Packet Service Branch, Telecom Australia, Melbourne, June.

Morriss, P. [1993]. "A Quantitative Approach to GIS Monitoring." Unpublished Honours thesis, Department of Surveying and Spatial Information Sciences, University of Tasmania, Hobart, Tasmania, Australia.

Myers, B.A. [1985]. "The Importance of Percent-done Progress Indicators for Computer-Human Interfaces", *CHI'85 Proceedings*, pp. 11-18 Association of Computing Machinery (ACM). April.

Newton, P.W., Zwart, P.R. and Crawford, J.R. with Cavill, M. [1990]. *High Speed Communication Networks and Industrial Graphics Systems: A Preliminary Survey of Potential Leading Edge Users of FASTPAC Services*, Contract Report commissioned by Fast Packet Services Group, Telecom Australia, Melbourne, Victoria, April.

Newton, P.W.and P.R. Zwart and M. Cavill [1992a]. "Inhibitors and Facilitators in High-Speed Networking of Spatial Information Systems." in *Networking Spatial Information Systems*, P.W. Newton, P.R. Zwart and M.E. Cavill, eds. Belhaven Press, London.

Newton, P.W.and P.R. Zwart and M. Cavill [1992b]. "Introducing the Next Wave: Networking Spatial Information Systems." in *Networking Spatial Information Systems,* Newton, P.W.and P.R. Zwart and M. Cavill, eds. Belhaven Press, London.

Palmer, D.W. [1984]. "A Land Information Network for New Brunswick". MScE. thesis, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B., Canada.

Petrosky, M. [1992]. "The Many Strands of Systems Management." , pp. 39-46. June.

Peuquet, D. and D. Marble [1990]. "PART IV GIS Internals — Data Representation and Analysis Techniques." *Introductory Readings in Geographic Information Systems*, D.J. Peuquet and D.F. Marble, eds., Taylor and Francis.

Pollock, R. and Y.C. Lee [1989]. "Landnet II Project Appendix B -- Review of Recent Technology". Appendix to contract report prepared for the Geographic Information Management Advisory Committee, Province of New Brunswick, Fredericton, N.B., Canada.

Popko, E. [1988]. "An Enterprise Implementation of AM/FM". Internal Paper, Geo-Facilities Information Systems Application Center, IBM Corporation, Houston, Texas, U.S.A.

Pretty, R. [1991]. "Fastpac - Nationwide High Speed Networking." *Australian Communications.* March.

Pretty, R.W. [1990]. "LANs, MANs and WANs." *Proceedings of the Australian Seminar on Networking Spatial Information Systems*, Telecom Australia Research Laboratories, Clayton, Victoria, Australia, June.

Pretty, R.W. [1992]. "LANs, MANs and WANs: Introductory Tutorial on Computer Networking". in *Networking Spatial Information Systems*, P.W. Newton, P.R. Zwart and M.E. Cavill, eds. Belhaven Press, London.

Rhind, D. [1992] "The Information Infrastructure of GIS." *Proceedings of the 5th International Symposium on Spatial Data Handling,* Vol. 1, pp. 1-19. Humanities and Social Sciences Computing Lab, University of South Carolina, Columbia, S.C., U.S.A.

Roberts, W.F. [1968]. "Some Basic Features of an Environmental Integrated Data Bank." *Proceedings of the Symposium on Land Registration and Data Banks,* Fredericton, New Brunswick, Canada, pp.30-33. (Reprinted in *The Canadian Surveyor,* Vol. XXIII, No. 1, March, 1969.)

Robinson, J. [1986]. *The Intelligent Infrastructure Report.* Available through A-E-C Automation Newsletter, Fountain Hills, Arizona, U.S.A.

Rosenblum, M. and J. Ousterhout [1992]. "The Design and Implementation of a Log-Structured File System". *ACM Transactions on Computer Systems*, February.

Sandberg, R. and D. Goldberg and S. Keiman and D. Walsh and B. Lyon [1985]. "Design and Implementation of the SUN Network Filesystem." *Proceedings of the USENIX Summer Conference*, pp. 119-130. June.

Sedunary, M.E. [1984]. "LOTS and the Nodal Approach to a Total Land Information System". *Proceedings of "The Decision Maker and Land Information Systems"--FIG Commission III International Symposium*, ed. A.C. Hamilton and J.D. McLaughlin. Edmonton, Alberta, Canada. October.

Sedunary, M.E. [1988]. "Land Information Systems – Reaping their Benefits." *Proceedings of the 1988 Annual Conference of the Urban and Regional Information Systems Association (URISA]*, Vol. 1, pp. 58-67.

Senter, R.J. [1969]. *Analysis of Data.* Scott, Foresman and Company, Glenview, Illinois, U.S.A.

Shoch, J.F. and J.A. Hupp. [1980]. "Measured performance of an Ethernet local network". *Communications of the ACM*, Vol. 23, No. 12, pp. 711-721. December.

Sinnreich, H. and J.F. Bottomley [1992]. "Any-to-Any Networking: Getting There From Here." *Data Communications*, pp. 69-80, September.

SLIC [1992]. *State Land Information Council Newsletter.* Special Issue on SLIC Initiatives. No. 15. New South Wales State Land Information Centre Directorate, Sydney, NSW, Australia. March.

Smith, G. [1992]. "WAN Internetworking — Emerging Standards and Services". *INTE NET*, January 1992, pp. 7-22.

Sloan, T.M and S. Dowers, B.M. Gittings, R.G. Healey and and T.C. Waugh [1992]. "Exploring GIS Performance Issues." *Proceedings of the 5th International Spatial Data Handling Symposium,* Vol. 1, pp. 154-165. Charleston, South Carolina, U.S.A.

Sondheim, M. and R. Menes [1991]. "General Framework for Geomatics Data Exchange." Proceedings of the GIS'91 Symposium, Forestry Canada FRDA Report 153, pp. 77-82, Vancouver, B.C. Canada.

Sproull. L. and S. Kiesler [1991]. *Connections: New Ways of Working in the Networked Organization.* MIT Press, Cambridge, MA., U.S.A.

Starling, R.N. and F. Dungca [1991]. "Guam Addresses On-Line Transaction Processing in a Land Information System - Titles, Permit Tracking and Taxation." Presentation at the 1991 NSW Land Information Management Conference, School of Surveying, University of New South Wales, Kensington, NSW, Australia. July.

Stefanovic, P. and J. Drummond [1987]. "Selection and Evaluation of Computer-Assisted Mapping and Geo Information Systems". *ITC Journal.* 1987-1

Stern, H. [1991]. *Managing NFS and NIS.* O'Reilly and Associates, California.

Stockton, B. [1990]. "Integrating GIS Within A Corporate Database Management Strategy." *Proceedings of the 1990 Annual Meeting of the Urban and Regional Information Systems Association (URISA].*, Vol. III, pp. 28-37. Urban and Regional Information Systems Association, Washington, D.C., U.S.A.

Stonebreaker, M. [1985]. "Tips on Benchmarking Data Base Systems." *Database Engineering*, Vol. 8, No. 1, pp. 10-18.

Stonebreaker, M. [1992]. An Overview of the Sequoia 2000 Project". *Proceedings of the CompCon'92 Conference*, San Francisco, California, U.S.A.

Sun Microsystems [1991]. *Sun Prestoserve Installation Manual.* Sun Microsystems Computer Corporation, Mountain View, California, U.S.A.

Telecom Australia [1992]. "FASTPAC: A Technical Description." *Technical brochure prepared by Fast Packet Services Group, Telecom Australia*, Melbourne, Victoria, Australia. March.

Teorey, T.J. and D.W. Bachmann, T.A. Makowski and M.M. Srivanivasan [1990]. "User Profile and Workload Analysis for Local Area Networks." Technical Report CITI-TR-90-3. Center for Information Technology Integration (CITI), University of Michigan, Ann Arbor, MI, US.A. June.

Tomlinson, R.F. [1967]. "An introduction to the geo-information system of the Canada Land Inventory." Department of Forestry and Rural Development, Government of Canada. Ottawa, Canada.

Tomlinson, R. and A.R. Boyle [1981]. "The State of Development of Systems for Handling Natural Resources Inventory Data". *Cartographica*, Vol. 18, No. 4, pp. 65-95.

Tonkin, B.A. and F.J.W. Symons [1992]. "Future Applications for MANs". in *Networking Spatial Information Systems*, P.W. Newton, P.R. Zwart and M.E. Cavill, eds. Belhaven Press, London.

Vermont OGIS [1989]. "Providing Access to the Vermont Geographic Information System: Legal Framework, Functional Access and Cost-Price Model for Sale of Products and Services." Appendix to contract report prepared by Plangraphics, Inc. for the Vermont Office of Geographic Information Services, Montpelier, VT., U.S.A.

Voelcker, John [1986]. "Helping Computers Communicate". *IEEE Spectrum*, March. pp.61-79.

Wagner, D. [1988]. "A Method of Evaluating Polygn Overlay Algorithms". *Proceedings of the 1988 ACSM-ASPRS Annual Convention*, Vol. 5, pp. 193-183.

Wagner, D.F. [1991]. "Development and Proof-of-Concept of a Comprehensive Performance Evaluation Methodology for Geographic Information Systems." Unpublished PhD. Dissertation. Department of Geography, Ohio State University, Columbus, OH, U.S.A. September.

Wagner, D.F. [1992]. Synthetic Test Design of Systematic Evaluation of Geographical Information Systems Performance". *Proceedings of the 5th International Spatial Data Handling Symposium,* Vol. 1, pp. 166-177. Charleston, South Carolina, U.S.A.

Webster, C. [1988]. "Disaggregated GIS Architecture. Lessons from recent developments in multi-site database management systems." *International Journal of Geographical Information Systems*, vol. 2, no. 1, pp. 67-80. January-March.

Whitehouse, B. [1993]. Personal communication with Operations Manager, Atlantic Centre for Remote Sensing of the Oceans, Halifax, Nova Scotia, Canada. June.

Wittie, L.D. [1991]. "Computer Networks and Distributed Systems". *IEEE Computer*, pp. 67-76. September.

Yang, W-P. [1992]. "A New Range Searching Algorithm for Large Point Databases". Unpublished MSc Thesis, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B., Canada.

Yang, O.W.W. and X-X Yao and K.M.S. Murthy [1992]. "Modelling and Performance Analysis of File Transfer in a Satellite Wide Area Network." *IEEE Journal on Selected Areas in Communications*, Vol. 10, No. 2, pp. 428-436.

Zhou, Q. [1991] Posting to the GIS-L Discussion List on Internet.

Zwart, P.R. [1991]. "Development of Spatial Information Standards." *Search.* Vol 22, No. 2. March.

Zwart, P.R. and D.J. Coleman [1991] "Introduction to Geographical Information Systems." *Presented at the Longman's Australian Conference on GIS*,. Melbourne, Victoria, Australia, July.

Zwart, P.R. and D.J. Coleman [1992]. "Tradeoffs Towards Real-time Spatial Information Networks." in *Networking Spatial Information Systems*, P.W. Newton, P.R. Zwart and M.E. Cavill, eds. Belhaven Press, London.

APPENDIX A

Selected Examples of System Usage Summaries from Participating Organisations

GIS Log Files – Tasmanian Forestry Commission

Users

Four log files were obtained from the Forestry Commission in early 1991. The files were workspace logs obtained from people deemed to be "representative" GIS users involved in mapping, field-office enquiries, forest planning and system management respectively. Details are included below:

Log File Name	User Name	Time Span *(dd/mm/yy)*
GIS	Bruce Haywood (GIS Manager)	02/02/91 – 15/02/91
Plan	Tom Kelly (Forest Planner)	15/05/89 – 19/02/91
Field	Phil Donnelly (Regional Forester)	20/12/88 – 02/11/90
Map	David Boden (Mapping Technician)	28/08/90 – 20/11/90

Indicators Employed

Log files were processed and compared using several combinations of indicators in order to best determine and compare representative usage by different groups. Preliminary summaries characterised GIS usage by noting the presence or absence of commands within a particular category, and then summarising its relative degree of usage in terms of both *Percentage Invoked* and *Percentage Connect*, where:

$$\text{Percentage Invoked} = \left[\frac{\text{No. of Times This Particular Command is Invoked}}{\text{Total No. of Commands Invoked}}\right] \times 100$$

$$\text{Percentage Connect} = \left[\frac{\text{Connect-Time Devoted to this Operation or Command}}{\text{Total Amount of Connect Time Overall}}\right] \times 100$$

The attached charts compare these two indicators for each user over the range of ArcInfo commands employed.

Qualifiers

- The research indicated that – while useful in summarising usage – the log entries may contain significant errors due to unforeseen bugs in the software itself. These bugs would occasionally cause the erroneous logging of very high *negative* execution times for (especially) ArcEdit, ArcPlot and INFO operations. ESRI was informed of this problem and have advised that it will be addressed in one of the Rev. 6.0x versions.
- Log files do not record the operations or commands invoked while working inside any of the other modules (e.g., ArcEdit, ArcPlot, TIN, etc.). Log entries only indicate the time each module was started and the amount of time spent therein.
- Elapsed execution times recorded in the log files include "stare time" – i.e., any time spent considering the next command or operation before actually invoking it. This is especially evident in ArcEdit and ArcPlot sessions.

Tas. Forestry Commission -- PLAN.Log Summary

Summaries										
Command	**Avg. date**	**Std. Dev.**	**Conn.**	**CPU**	**I/O**	**Times**	**%**	**%**	**%**	**%**
						Invoked	**Conn.**	**CPU**	**I/O**	**Invoked**
ADDITEM			33	679	226	134	0%	1%	1%	6%
APPEND			1	20	2	1	0%	0%	0%	0%
ARCDXF			2	90	0	1	0%	0%	0%	0%
ARCEDIT			920	3897	2126	188	10%	4%	10%	8%
ARCPLOT			3005	17193	3487	379	34%	20%	17%	17%
ATTEDIT			838	2497	1594	63	9%	3%	8%	3%
BUILD			27	622	54	21	0%	1%	0%	1%
CLEAN			1183	23262	6863	102	13%	27%	33%	5%
COPY			77	917	1055	129	1%	1%	5%	6%
CREATELABELS			47	718	346	121	1%	1%	2%	5%
DISSOLVE			124	4109	738	102	1%	5%	4%	5%
DROPITEM			19	626	19	71	0%	1%	0%	3%
EXTERNALALL			5	126	4	3	0%	0%	0%	0%
GENERATE			15	111	17	13	0%	0%	0%	1%
HPGL			0	2	3	1	0%	0%	0%	0%
IDENTITY			187	4766	325	8	2%	5%	2%	0%
INFO			1171	5957	1282	300	13%	7%	6%	13%
INTERSECT			111	4142	118	18	1%	5%	1%	1%
KILL			100	1518	867	420	1%	2%	4%	19%
LIBRARIAN			5	61	7	2	0%	0%	0%	0%
LIST			0	1	0	1	0%	0%	0%	0%
MAPJOIN			86	1828	172	6	1%	2%	1%	0%
NODEERRORS			0	15	6	3	0%	0%	0%	0%
POLYGRID			2	18	1	1	0%	0%	0%	0%
POSTSCRIPT			17	622	4	2	0%	1%	0%	0%
RENAME			21	343	109	83	0%	0%	1%	4%
RESELECT			40	931	370	38	0%	1%	2%	2%
UNION			265	10464	1012	5	3%	12%	5%	0%
ZETA			541	1350	77	42	6%	2%	0%	2%
							100%	100%	100%	100%

Tas. Forestry Commission -- MAP.Log Summary

Summaries										
Command	**Avg. date**	**Std. Dev.**	**Conn.**	**CPU**	**I/O**	**Times**	**%**	**%**	**%**	**%**
						Invoked	**Conn.**	**CPU**	**I/O**	**Invoked**
ADDITEM			1	34	1	6	0%	0%	0%	1%
ARCEDIT			8644	33241	4307	221	42%	37%	37%	27%
ARCFONT			1	26	3	7	0%	0%	0%	1%
ARCPLOT			3991	7843	670	82	20%	9%	6%	10%
BUILD			155	5504	317	12	1%	6%	3%	1%
CLEAN			386	13195	971	10	2%	14%	8%	1%
CLIP			30	1241	325	1	0%	1%	3%	0%
COPY			12	176	290	25	0%	0%	2%	3%
DISSOLVE			4	69	20	2	0%	0%	0%	0%
DROPITEM			29	1137	75	51	0%	1%	1%	6%
ELIMINATE			9	281	119	4	0%	0%	1%	0%
ERASE			24	937	235	1	0%	1%	2%	0%
EXPORT			66	2956	31	10	0%	3%	0%	1%
FONTCREATE			1	10	2	3	0%	0%	0%	0%
GENERATE			7	17	8	2	0%	0%	0%	0%
IDEDIT			0	8	1	1	0%	0%	0%	0%
IDENTITY			282	9868	2096	5	1%	11%	18%	1%
IMPORT			96	1574	7	2	0%	2%	0%	0%
INFO			6043	5506	700	192	30%	6%	6%	23%
KILL			9	207	155	73	0%	0%	1%	9%
LABELERRORS			2	38	5	22	0%	0%	0%	3%
LIBRARIAN			13	91	18	2	0%	0%	0%	0%
MARKEREDIT			30	49	12	5	0%	0%	0%	1%
NODEERRORS			9	256	42	15	0%	0%	0%	2%
OURBATCH			12	42	10	15	0%	0%	0%	2%
RENAME			31	222	44	17	0%	0%	0%	2%
RESELECT			15	348	48	5	0%	0%	0%	1%
RESTOREARCEDIT			1	48	5	2	0%	0%	0%	0%
ROTATEPLOT			21	646	12	8	0%	1%	0%	1%
UNION			114	4431	1177	2	1%	5%	10%	0%
ZETA			398	1025	64	20	2%	1%	1%	2%
							1	1	1	1

Explanatory Notes

GIS Log Files – Victorian Department of Conservation and Environment

Users

The log files of six Arc/Info users were obtained from the Department in early 1991. With the exception of *Fauna* and *LIM*, the remaining files were workspace logs obtained from people deemed to be "representative" GIS users involved involved in – for example – mapping, forest planning and systems management/analysis. Details are included below:

Log File Name	User Name	Time Span *(dd/mm/yy)*
Fauna	FAUNA Database Users	08/05/90 – 28/02/91
LIM	LIM Database Users	07/11/90 – 13/03/91
Reg_1	Regional Office Staff Member	22/01/90 – 13/03/91
Reg_2	" " " "	13/02/90 – 06/03/91
Anal_1	GIS Analyst - Head Office	31/07/90 – 23/01/91
Anal_2	" " " "	12/05/89 – 01/02/91

Indicators Employed

Log files were processed and compared using several combinations of indicators in order to best determine and compare representative usage by different groups. Preliminary summaries characterised GIS usage by noting the presence or absence of commands within a particular category, and then summarising its relative degree of usage in terms of both *Percentage Invoked* and *Percentage Connect*, where:

$$\text{Percentage Invoked} = \left[\frac{\text{No. of Times This Particular Command is Invoked}}{\text{Total No. of Commands Invoked}}\right] \times 100$$

$$\text{Percentage Connect} = \left[\frac{\text{Connect-Time Devoted to this Operation or Command}}{\text{Total Amount of Connect Time Overall}}\right] \times 100$$

The attached charts compare these two indicators for each user over the range of ArcInfo commands employed.

Qualifiers

- The research indicated that – while useful in summarising usage – the log entries may contain significant errors due to unforeseen bugs in the software itself. These bugs would occasionally cause the erroneous logging of very high *negative* execution times for (especially) ArcEdit, ArcPlot and INFO operations. ESRI was informed of this problem and have advised that it will be addressed in one of the Rev. 6.0x versions.

- Log files do not record the operations or commands invoked while working inside any of the other modules (e.g., ArcEdit, ArcPlot, TIN, etc.). Log entries only indicate the time each module was started and the amount of time spent therein.

- Elapsed execution times recorded in the log files include "stare time" – i.e., any time spent considering the next command or operation before actually invoking it. This is especially evident in ArcEdit and ArcPlot sessions.

VIC C+E GIS_Analyst.2 LOG SUMMARY

Summaries								
Command	**Conn.**	**CPU**	**I/O**	**Times Invoked**	**% Conn.**	**% CPU**	**% I/O**	**% Invoked**
additem	5	112	0	32	0%	0%	0%	3%
ap / arcplot	478	951	88	70	20%	3%	3%	7%
ae / arcedit	779	1641	112	43	32%	6%	4%	5%
buffer								
build	1	56	3	3	0%	0%	0%	0%
clean	22	666	29	16	1%	2%	1%	2%
clip	3	1407	245	3	0%	5%	9%	0%
copy	6	110	99	18	0%	0%	4%	2%
copyinfo	6	118	21	90	0%	0%	1%	9%
create	0	6	0	3	0%	0%	0%	0%
createlabels	1	3	0	2	0%	0%	0%	0%
cw	0							
dissolve	2	42	1	2	0%	0%	0%	0%
dropitem	25	865	1	236	1%	3%	0%	25%
export	89	1299	22	1	4%	5%	1%	0%
frequency	65	1671	9	15	3%	6%	0%	2%
HPGL	0							
idedit	1	2	1	2	0%	0%	0%	0%
identity	367	13005	1919	7	15%	48%	69%	1%
INFO	262	2999	76	221	11%	11%	3%	23%
intersect	0							
joinitem	1	14	0	4	0%	0%	0%	0%
kill	8	115	62	70	0%	0%	2%	7%
librarian	5	27	2	1	0%	0%	0%	0%
list	4	3	1	8	0%	0%	0%	1%
mapjoin					0%	0%	0%	0%
project	56	1317	41	11	2%	5%	1%	1%
pullitems	1	8	0	4	0%	0%	0%	0%
rebox	1	3	0	4	0%	0%	0%	0%
rename	0	8	0	5	0%	0%	0%	1%
reselect	11	247	11	16	0%	1%	0%	2%
rotateplot					0%	0%	0%	0%
tables	200	526	35	64	8%	2%	1%	7%
transform	1	7	2	3	0%	0%	0%	0%
Totals	2400	27228	2780	954	1	1	1	

VIC C+E FAUNA DATABASE LOG SUMMARY

Summaries								
Command	**Conn.**	**CPU**	**I/O**	**Times**	**%**	**%**	**%**	**%**
				Invoked	**Conn.**	**CPU**	**I/O**	**Invoked**
additem					0%	0%	0%	0%
ap / arcplot	1741	9818	1670	97	99%	98%	98%	85%
ae / arcedit								
buffer								
build								
clean								
clip								
copy								
copyinfo								
create								
createlabels								
cw								
dissolve	15	181	21	7	1%	2%	1%	6%
dropitem								
export								
frequency								
HPGL	11	47	2	3	1%	0%	0%	3%
idedit								
identity								
INFO								
intersect								
joinitem								
kill	0	11	3	7	0%	0%	0%	6%
librarian								
list								
mapjoin								
project								
pullitems								
rebox								
rename								
reselect								
rotateplot								
tables								
transform								
Totals	**1767**	**10057**	**1696**	**114**	**1**	**1**	**1**	**1**

Explanatory Notes

RIOS Usage Log Files – Sydney Water Board

Users

Six RIOS log files were obtained from the Sydney Water Board covering sessions during September 1991. These files were workspace logs obtained from people deemed to be "representative" GIS users involved involved in – for example – drafting, engineering, business office enquiries and systems management/analysis. Details are included below:

Log File Name	Type of User	Time Span
LAR	Engineering & Systems	Sept. 12th Sept. 30th, 1991
LCM	Business Office	Sept. 13th Sept. 27th, 1991
SQZ	Drafting	Sept. 13th Sept. 24th, 1991
HOO	Business Office	Sept. 27th, 1991
HUM		Sept. 13th Sept. 24th, 1991
POV		Sept. 12th Sept. 26th, 1991

Indicators Employed

Preliminary summaries characterised RIOS usage by noting the presence or absence of commands within a particular category, and then summarising its relative degree of usage in terms of *Percentage Invoked,* where:

$$\text{Percentage Invoked} = \left[\frac{\text{No. of Times This Particular Command is Invoked}}{\text{Total No. of Commands Invoked}} \right] \times 100$$

The attached charts compare these two indicators for each user over the range of ArcInfo commands employed.

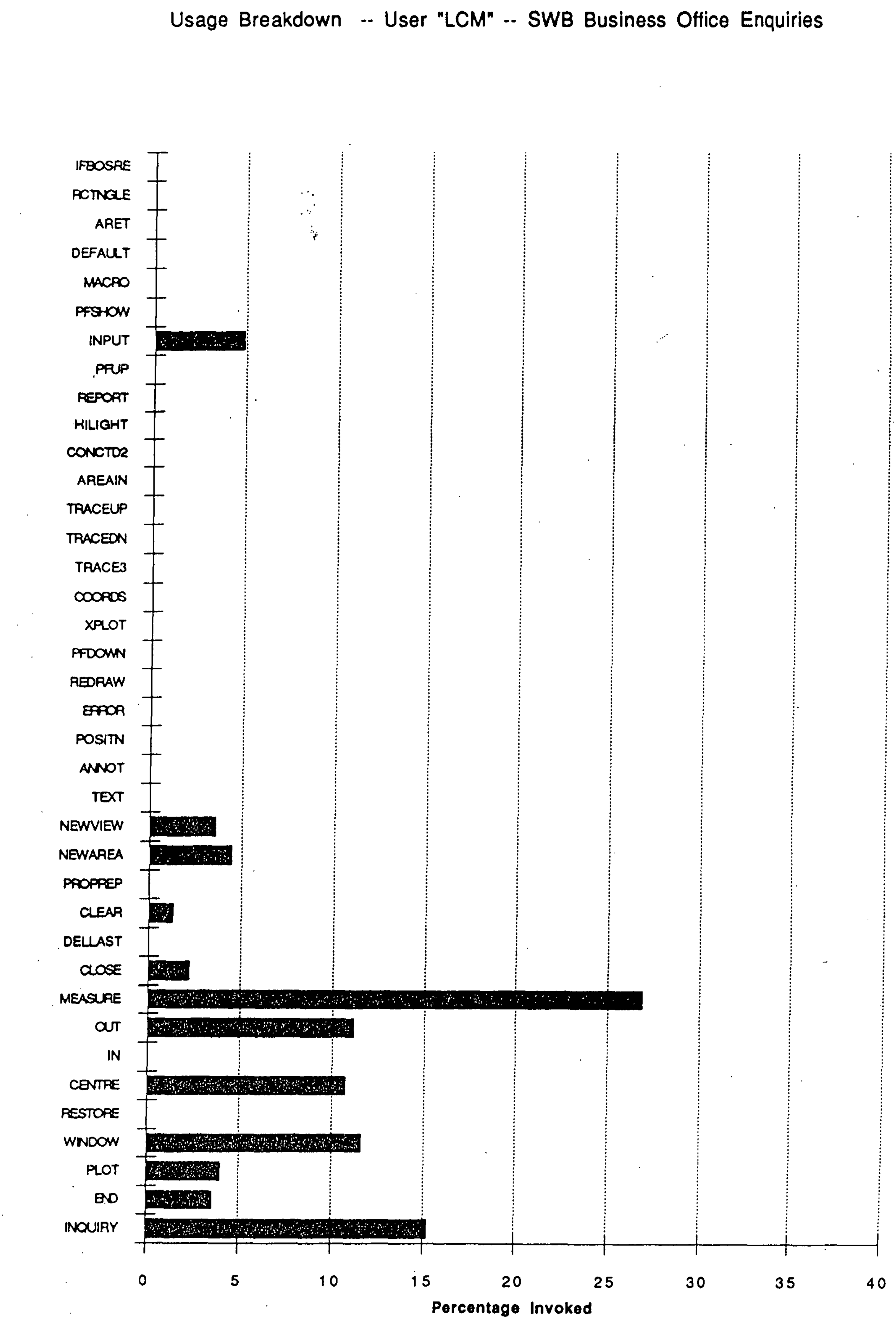
NOTE: After consultation with Greg Cox at the Water Board, RIOS commands were broadly separated into four categories -- Session Control, Graphics and Text Addition & Enhancement, Routine Enquiries and Special or Complex Enquiries. A second set of charts was then prepared which illustrated the same figures broken into these categories.

Qualifiers

- These charts only indicate the relative numbers of times each command is invoked. Conclusions drawn from such a summary on its own may be misleading and must be considered only in concert with other factors. Experience with log files on other systems indicates there is not necessarily any correlation between the number of times a command is invoked and the overall proportion of elapsed time spent completing such operations. Indeed, a small percentage of commands may consume most of the system time. Given the time constraints on our own research, we were unable to extract, process and interpret execution-time data for each command during the first round of processing.

BREAKDOWN OF WATER BOARD RIOS COMMANDS

	PF Key	Session Control	Routine Enquiry	Hardcopy Plots	Special Enquiry	Graphics Add'n.	Other	Type No.
COMMAND								
INQUIRY	2		✔					2
END	3	✔						1
PLOT	4			✔				3
WINDOW	5		✔					2
RESTORE	6		✔					2
CENTRE (or PAN)	7		✔					2
IN	8		✔					2
OUT	9		✔					2
MEASURE	10				✔			4
CLOSE	11				✔			4
DELLAST	12				✔			4
CLEAR	13				✔			4
PROPREP	14		✔					2
NEWAREA	16		✔					2
NEWVIEW	17		✔					2
TEXT	18					✔		5
ANNOT	19					✔		5
POSITN	20		✔					2
ERROR	21						✔	6
REDRAW	22		✔					2
PFDOWN	24		✔					2
XPLOT							✔	6
COORDS			✔					2
TRACE3					✔			4
TRACEDN					✔			4
TRACEUP					✔			4
AREAIN					✔			4
CONCTD2							✔	6
HILIGHT					✔			4
REPORT			✔					2
PFUP			✔					2
INPUT					✔			4
PFSHOW		✔						1
MACRO							✔	6
DEFAULT		✔						1
ARET					✔			4
RCTNGLE					✔			4
IFBOSRE							✔	6
BUFFER								4
DYVIEW								4
FREEND								4
FILL								4
ISPF								1
RCTNGLD								4
STORE								4
XING								4


Usage Breakdown -- User "LCM" -- SWB Business Office Enquiries
IFBOSRE
RCTNGLE
ARET
DEFAULT
MACRO
PFSHOW
INPUT
PFUP
REPORT
HILIGHT
CONCTD2
AREAIN
TRACEUP
TRACEDN
TRACE3
COORDS
XPLOT
PFDOWN
REDRAW
ERROR
POSITN
ANNOT
TEXT
NEWVIEW
NEWAREA
PROPREP
CLEAR
DELLAST
CLOSE
MEASURE
OUT
IN
CENTRE
RESTORE
WINDOW
PLOT
END
INQUIRY
0
5
10
15
20
25
30
35
40
Percentage Invoked

APPENDIX B

Examples of Command Scripts of Major GIS Performance Tests

Appendix B

EXAMPLES OF COMMAND SCRIPTS OF MAJOR GIS PERFORMANCE TESTS

(written in Arc Macro Language for the ESRI Arc/Info Software)

MIXA.aml *(Single-User Test run on UNIX Server)*

```
&severity &error &ignore
&severity &warning &ignore
&station 9999
&echo &on
&pt &on
/*
/*
&watch fp_mixa2
/*          Kill any unwanted coverages previously generated
kill /beta/b50g1
kill /beta/b50u1
kill /beta/b50c1
kill /alpha/a50g1
kill /alpha/a50u1
kill /alpha/a50c1
kill /beta/p50g1
kill /gisdat/p50c1
/*
/*          External any required files from other directories
external /gisdat/p50g
external /gisdat/p50g11
external /gisdat/p50d1
external /gisdat/p50d11
external /gisdat/p50u11
external /gisdat/port
external /beta/b50g
external /beta/b50g11
external /beta/b50d1
external /beta/b50d11
external /beta/b50u11
external /beta/port
/*
/*
&do I := 1 &to 20
/*
   &type Now executing Copy BAA  # %I%
   copy /beta/b50u11 /gisdat/p50c1
/*
   &type Now executing Clean BAA  # %I%
   clean /beta/b50g /gisdat/p50g1 0 0.002 poly
     kill /gisdat/p50g1
/*
   &type Now executing Copy AAAw  # %I%
   copy /gisdat/p50u11 /alpha/a50c1
     kill /alpha/a50c1

/*
   arcedit
       &type Now executing ArcEdit Draw # %I%
        symbolset color
       editcoverage /gisdat/p50d11
        mapextent /gisdat/p50d11
        drawenvironment all on
```

MIXA.aml (cont'd)

```
        draw
        clear
    quit
/*
    &type Now executing Clean AAAw  # %I%
    clean /gisdat/p50g /alpha/a50g1 0 0.002 poly
      kill /alpha/a50g1
/*
    arcplot
        &type Now executing Arcplot Session # %I%
        symbolset color
        mapwarp /gisdat/port
        mape image /gisdat/port
        image /gisdat/port
        clear
    quit
/*
    &type Now executing Copy AAB # %I%
    copy /gisdat/p50c1 /beta/b50c11
      kill /beta/b50c11
/*
    arcedit
        &type Now executing ArcEdit Draw # %I%
        symbolset color
        editcoverage /beta/b50d11
        mapextent /beta/b50d11
        drawenvironment all on
        draw
        clear
    quit
/*
    &type Now executing Clean AAB  # %I%
    clean /gisdat/p50g /beta/p50g1 0 0.002 poly
      kill /beta/p50g1
/*
    arcplot
        &type Now executing Arcplot Session # %I%
        symbolset color
        mapwarp /beta/port
        mape image /beta/port
        image /beta/port
        clear
    quit
/*
    &type Now executing Copy BAB  # %I%
    copy /beta/b50u11 /beta/b50c1
      kill /beta/b50c1
/*
    &type Now executing Clean BAB  # %I%
    clean /beta/b50g /beta/b50g1 0 0.002 poly
      kill /beta/b50g1
/*
      kill /gisdat/p50c1
/*
&end
&watch &off
/*
/*
&pt &off
&return
```

MIXB aml *(Single-User Test run on remote client workstation)*

```
&severity &error &ignore
&severity &warning &ignore
&station 9999
&echo &on
&pt &on
/*
/*
&watch fp_mixb2
/*          Kill any unwanted coverages previously generated
kill /beta/b50g1
kill /beta/b50u1
kill /beta/b50c1
kill /alpha/a50g1
kill /alpha/a50u1
kill /alpha/a50c1
kill /beta/p50g1
/*
/*          External any required files from other directories
external /gisdat/p50g
external /gisdat/p50g11
external /gisdat/p50d1
external /gisdat/p50d11
external /gisdat/p50u11
external /gisdat/port
external /alpha/b50g
external /alpha/b50g11
external /alpha/b50d1
external /alpha/b50d11
external /alpha/b50u11
external /alpha/port
/*
/*
&do I := 1 &to 20
/*
   &type Now executing Copy ABA  # %I%
   copy /gisdat/p50u11 /gisdat/p50c1
     kill /gisdat/p50c1
/*
   &type Now executing Clean ABA  # %I%
   clean /gisdat/p50g /gisdat/p50g1 0 0.002 poly
     kill /gisdat/p50g1
/*
   &type Now executing Copy BBB  # %I%
   copy /beta/b50u11 /beta/b50c1
     kill b50c1
/*
   arcedit
       &type Now executing ArcEdit Draw # %I%
        symbolset color
       editcoverage /gisdat/p50d11
        mapextent /gisdat/p50d11
        drawenvironment all on
        draw
        clear
   quit
/*
   &type Now executing Clean BBB  # %I%
   clean /beta/b50g /beta/b50g1 0 0.002 poly
     kill b50g1
/*
```

MIXB.aml (cont'd.)

```
    arcplot
        &type Now executing Arcplot Session # %I%
         symbolset color
        mapwarp /gisdat/port
        mape image /gisdat/port
        image /gisdat/port
        clear
    quit
/*
    &type Now executing Copy BBA # %I%
    copy /beta/b50u11 /gisdat/p50c1
      kill /gisdat/p50c1
/*
    arcedit
        &type Now executing ArcEdit Draw # %I%
         symbolset color
        editcoverage /beta/b50d11
        mapextent /beta/b50d11
        drawenvironment all on
        draw
        clear
    quit
/*
    &type Now executing Clean BBA # %I%
    clean /beta/b50g /gisdat/p50g1 0 0.002 poly
      kill /gisdat/p50g1
/*
    arcplot
        &type Now executing Arcplot Session # %I%
         symbolset color
        mapwarp /beta/port
        mape image /beta/port
        image /beta/port
        clear
    quit
/*
    &type Now executing Copy ABB # %I%
    copy /gisdat/p50u11 /beta/b50c1
      kill /beta/b50c1
/*
    &type Now executing Clean ABB # %I%
    clean /gisdat/p50g /beta/b50g1 0 0.002 poly
      kill /beta/b50g1
/*
&end
&watch &off
/*
/*
&pt &off
&return
```

*****

SNOISE.aml ***(Designed to create heavy load on Server memory and disk)***

```
&severity &error &ignore
&severity &warning &ignore
&station 9999
&echo &on
&pt &on
/*
/*
&watch ln_noise
/*          Kill any unwanted coverages previously generated
kill a50c1
kill a50c2
kill a50c3
kill a50c4
/*
/*          External any required files from other directories
external /gisdat/p50g
external /gisdat/p50g12
external /gisdat/p50d2
external /gisdat/p50d12
external /gisdat/p50u12
external /gisdat/port
/*
/*
&do I := 1 &to 60
/*
   copy /gisdat/p50u12 a50c1
   kill a50c1

/*
   copy /gisdat/p50d12 a50c2
     kill a50c2
/*
   copy /gisdat/p50d2 a50c3
     kill  a50c3
/*
   copy /gisdat/p50g12 a50c4
     kill a50c4
/*
/*
&end
&watch &off
/*
/*
```

***

SNOISE2.aml ***(Designed to create medium load on Server memory and disk)***

```
&severity &error &ignore
&severity &warning &ignore
&station 9999
/*             Kill any unwanted coverages previously generated
kill a50c1
kill a50c2
kill a50c3
kill a50c4
/*
/*             External any required files from other directories
external /gisdat/p50g
external /gisdat/p50g12
external /gisdat/p50d2
external /gisdat/p50d12
external /gisdat/p50u12
external /gisdat/port
/*
/*
&do I := 1 &to 60
/*
   arcedit
       &type Now executing ArcEdit Draw # %I%
        symbolset color
       editcoverage /gisdat/p50d12
        mapextent /gisdat/p50d12
        drawenvironment all on
        draw
        clear
   quit
/*
   copy /gisdat/p50u12 a50c1
   kill a50c1
/*
   arcedit
       &type Now executing ArcEdit Draw # %I%
        symbolset color
       editcoverage /gisdat/p50u12
        mapextent /gisdat/p50u12
        drawenvironment all on
        draw
        clear
   quit
/*
   copy /gisdat/p50d12 a50c2
     kill a50c2
/*
/*
   arcplot
       &type Now executing Arcplot Session # %I%
        symbolset color
        mapwarp /gisdat/bport
        mape image /gisdat/bport
        image /gisdat/bport
        clear
   quit
/*
&end
/*
/*
&return
```

TRAFFTEST.aml ***(Designed to create defined load on across network and on client and server disks)***

```
&severity &error &ignore
&severity &warning &ignore
&station 9999
&echo &on
&pt &on
/*
/*
&watch ln_traff000
/*          Kill any unwanted coverages previously generated
kill a50g1
kill a50u1
kill a50c1
kill /beta/c50g1
kill /beta/c50u11
/*
/*          External any required files from other directories
external /gisdat/p50u11
external /gisdat/port
/*
&do I := 1 &to 15
/*
   &type Now executing Copy ABB # %I%
   copy /gisdat/p50u11 /beta/c50u11
     kill /beta/c50u11
/*
   arcplot
       &type Now executing Arcplot Session # %I%
       symbolset color
       mapwarp /gisdat/port
       mape image /gisdat/port
       image /gisdat/port
       clear
   quit
/*
   &type Now executing Copy BBA # %I%
   copy /beta/b50u11 /gisdat/b50u11
     kill /gisdat/b50u11
/*
   &type Now executing Copy ABA # %I%
   copy /gisdat/p50u11 /gisdat/c50u11
     kill /gisdat/c50u11
/*
&end
&watch &off
/*
/*
&echo &off
&pt &off
&return
```

APPENDIX C

Results of FASTPAC/GIS Performance Testing

C.1 UNIX File Transfer and Copying Tests (1991)

C.2 GIS Operations (1991 — Single User)

C.3 GIS Operations (1992 — Single User)

C.4 GIS Operations (1992 — Varying Traffic and Server Load)

		MEAN EXECUTION TIMES (secs)				MEAN TRANSFER RATES (kbits/sec)					STANDARD DEVIATIONS		
		Stand-Alone	LAN	FASTPAC	FASTPAC	Stand-Alone	LAN	FASTPAC	FASTPAC	Stand-Alone	LAN	FASTPAC	FASTPAC
OPERATIONS	**File Size**			(Short)	(Long)			(Short)	(Long)			(Short)	(Long)
	(Kb)												
cp (A to A internal)	3559.00	9.43				3020.37				0.49			
(Sever to Server)	9793.00	21.75				3602.58				0.34			
cp (B to B internal)	3559.00	10.19				2795.03				0.76			
(Client to Client)	9793.00	31.44				2491.86				0.47			
cp (A to B)	3559.00		49.65	50.36	52.20		573.49	565.37	545.44		0.43	0.94	0.62
(Sever to Client)	9793.00		146.41	149.15	159.06		535.09	525.27	492.54		0.74	1.15	0.64
cp (B to A)	3559.00		41.60	43.08	43.22		684.42	660.91	658.77		0.75	1.36	1.13
(Client to Server)	9793.00		124.71	125.38	125.93		628.19	624.85	622.12		0.50	1.49	0.63
		LAN	**FP Short**	**FP Long**	**TRL**					**LAN**	**FP Short**	**FP Long**	**TRL**
ftp (User @Server)	9793.00												
GET from Client		78.20	78.60	78.80	79.20		1001.84	996.74	994.21	0.45	0.55	0.45	0.84
PUT to Client		77.00	79.40	80.20	80.20		1017.45	986.70	976.86	0.71	0.55	0.45	0.45
ftp (User@Client)	9793.00												
GET from Server		74.80	75.20	77.20	n/a		1047.38	1041.81	1014.82	0.45	0.84	0.45	n/a
PUT to Server		78.40	78.40	80.00	n/a		999.29	999.29	979.30	0.55	0.89	0.71	n/a

		MEAN (sec)				**STD. DEV.**			
		Stand-Alone				Stand-Alone			
Configuration	**Operation**								
AAA	clean (A-A-A)	53.27				0.88			
AAA	copy (A-A-A)	34.14				0.66			
AAA	draw (A to A)	33.33				0.49			
AAA	polygons (A to A)	4.33				0.49			
AAA	polygonshades (A to A)	6.93				0.70			
AAA	reselect (A to A)	1.47				0.52			
AAA	union (A-A-A)	292.07				1.75			
•••									
				MEAN (secs)				**STD.DEV.**	
		TRL LAN	FPAC LAN	FASTPAC	FASTPAC	TRL LAN	FPAC LAN	FASTPAC	FASTPAC
				(Short)	(Long)			(Short)	(Long)
AAB	clean (A-A-B)		69.80		77.40		0.84		1.14
AAB	copy (B-A-B)		53.40		54.80		3.13		1.79
AAB	copy (B-A-A)		21.60		23.20		0.55		0.45
AAB	copy (A-A-B)		52.80		54.40		0.84		0.55
AAB	draw (B to A)		33.00		34.00		0.00		1.22
AAB	polygons (B to A)		3.80		4.60		0.45		0.55
AAB	polygonshades (B to A)		6.80		6.80		0.45		0.84
AAB	reselect (B to A)		1.20		1.60		0.45		0.55
AAB	union (B-A-B)		431.20		472.40		1.10		3.78
AAB	union (B-A-A)		295.10		299.20		3.07		3.42
BAA	clean (A-B-A)	68.27	67.80	75.50	78.30	0.80	1.87	0.93	1.49
BAA	copy (A-B-A)	51.33	47.80	52.25	53.00	0.82	0.79	0.71	1.41
BAA	copy (A-B-B)	22.73	24.30	26.13	27.30	0.80	2.50	0.83	0.82
BAA	copy (B-B-A)	45.60	42.50	43.88	46.30	0.74	0.85	0.64	1.34
BAA	draw (A to B)	33.07	32.60	34.00	34.20	0.88	0.52	0.76	0.42
BAA	polygons (A to B)	3.80	4.10	4.38	4.40	0.41	0.32	0.52	0.52
BAA	polygonshades (A to B)	6.20	6.00	6.38	6.70	0.43	0.00	0.52	0.48
BAA	reselect (A to B)	1.20	1.10	1.63	1.50	0.41	0.32	0.52	0.53
BAA	union (A-B-A)	419.47	398.60	436.25	446.30	2.36	1.35	2.12	4.03
BAA	union (A-B-B)	310.33	315.60	325.00	322.11	1.18	1.84	3.54	1.76

				MEAN				STD.DEV.	
		TRL LAN	FPAC LAN	FASTPAC	FASTPAC	Stand-Alone	LAN	FASTPAC	FASTPAC
Configuration	**Operation**			(Short)	(Long)			(Short)	(Long)
BAB	clean (B-B-B)	57.07	56.83	60.50	60.40	0.70	0.41	0.53	1.17
BAB	copy (B-B-A)	47.67	44.00	48.00	48.40	0.82	0.89	2.11	1.07
BAB	copy (B-B-B)	26.53	31.00	35.70	30.33	1.77	3.03	2.36	2.78
BAB	draw (B to B)	35.13	35.00	35.56	35.60	0.92	1.26	0.88	1.26
BAB	polygons (B to B)	5.20	6.00	5.30	5.10	0.68	0.89	0.82	0.57
BAB	polygonshades (B to B)	6.00	6.50	6.50	6.80	0.00	0.55	0.53	0.42
BAB	reselect (B to B)	1.33	1.17	1.30	1.30	0.49	0.41	0.48	0.48
BAB	union (B-B-B)	308.13	310.50	322.67	322.11	1.77	2.35	3.16	4.08
CAA	clean (A-C-A)	73.38	73.07	88.89	93.20	1.04	1.16	0.60	0.63
CAA	copy (A-C-B)	44.46	49.27	50.56	50.90	0.66	0.96	0.53	0.57
CAA	copy (A-C-A)	52.64	51.67	58.13	58.70	1.01	0.82	0.99	1.16
CAA	copy (C-C-C)	53.15	52.29	56.78	57.60	1.68	0.73	1.39	1.17
CAA	draw (A to C)	34.79	34.67	38.63	39.00	0.70	0.62	0.52	0.94
CAA	polygons (A to C)	4.00	4.00	4.67	4.80	0.28	0.00	0.50	0.79
CAA	polygonshades (A to C)	6.62	6.40	7.00	6.90	0.51	0.51	0.00	0.32
CAA	reselect (A to C)	1.31	1.64	1.78	2.00	0.48	0.63	0.44	0.00
CAA	union (A-C-A)	415.38	414.80	464.00	475.00	2.43	2.34	2.06	1.94
CAA	union (A-C-C)	414.64	411.27	461.13	468.90	2.21	2.40	1.89	4.04
CAB	clean (B-C-B)	72.93	73.80	79.60	81.10	0.88	1.15	0.52	0.57
CAB	copy (B-C-B)	46.00	50.80	52.20	52.70	0.93	1.32	0.63	2.00
CAB	copy (B-C-A)	45.07	47.20	50.70	51.40	1.49	1.01	1.06	0.70
CAB	draw (B to C)	35.20	35.33	38.30	39.00	0.86	0.72	0.48	0.47
CAB	polygons (B to C)	3.60	3.60	3.30	3.20	0.51	0.51	0.48	0.42
CAB	polygonshades (B to C)	6.27	6.00	6.30	6.50	0.46	0.00	0.48	0.53
CAB	reselect (B to C)	1.13	1.20	1.60	1.70	0.35	0.41	0.52	0.48
CAB	union (B-C-B)	414.60	436.07	450.70	455.70	3.36	3.37	2.06	1.70
CAB	union (B-C-C)	405.93	407.87	453.40	463.20	2.40	1.68	1.78	2.20

APPENDIX C.3 -- 1992 LAN/FASTPAC Testing Results -- By Configuration

		MEAN (secs)				STD. DEVIATION (secs)			
		NFSD = 0		NFSD = 8		NFSD = 0		NFSD = 8	
		LAN	FASTPAC	LAN	FASTPAC	LAN	FASTPAC	LAN	FASTPAC
Config.	Operation								
AAA	clean AAA	45.50				0.71			
AAA	copy AAA	25.30				0.48			
AAA	draw (to A from A)	28.70				0.48			
AAA	image (to A from A)	7.40				0.52			
MIX_A	Clean AAAw	44.30	45.70	45.73	44.50	0.67	1.06	0.88	0.55
MIX_A	Clean AAB	67.20	87.30	62.87	74.67	0.63	0.67	1.06	0.52
MIX_A	Clean BAA	42.73	45.50	44.40	45.17	0.70	0.85	0.63	0.75
MIX_A	Clean BAB	70.20	87.20	62.33	76.60	1.01	0.79	0.90	0.55
MIX_A	Copy AAAw	24.70	26.00	32.00	32.67	0.67	0.00	0.38	0.52
MIX_A	Copy AAB	42.70	56.00	42.13	53.50	0.48	0.00	0.35	0.55
MIX_A	Copy BAA	13.20	22.00	12.60	14.50	0.68	0.67	0.51	0.55
MIX_A	Copy BAB	52.13	72.80	55.53	67.00	0.35	0.42	0.52	0.00
MIX_A	Draw (to A from A)	28.50	27.30	27.00	27.00	0.48	0.48	0.53	0.63
MIX_A	Draw (to A from B)	29.47	29.90	27.47	28.00	0.52	0.32	0.52	0.00
MIX_A	Image (to A from A)	7.60	8.00	7.93	7.50	0.52	0.82	0.70	0.84
MIX_A	Image (to A from B)	18.60	31.80	13.33	18.50	0.51	0.42	0.49	0.55
MIX_B	Clean ABA	68.40	89.00	66.00	81.80	1.17	0.67	0.93	1.87
MIX_B	Clean ABB	44.33	46.60	44.47	45.00	0.50	0.70	0.64	0.82
MIX_B	Clean BBA	66.80	87.40	68.40	84.90	1.03	1.43	0.74	0.32
MIX_B	Clean BBB	45.00	46.50	46.40	N/A	1.25	0.97	1.68	N/A
MIX_B	Copy ABA	48.10	73.10	51.07	61.40	0.32	0.32	0.46	0.52
MIX_B	Copy ABB	13.50	21.90	12.93	15.20	0.53	0.57	0.59	0.63
MIX_B	Copy BBA	39.90	57.10	37.93	47.80	0.32	0.32	0.26	0.42
MIX_B	Copy BBB	25.30	26.20	26.33	N/A	0.48	0.42	0.62	N/A
MIX_B	Draw (to B from A)	29.30	29.50	28.87	29.60	0.67	0.71	0.64	0.52
MIX_B	Draw (to B from B)	26.40	26.80	28.67	N/A	0.52	0.42	0.62	N/A
MIX_B	Image (to B from A)	19.00	31.30	13.27	18.50	0.00	0.48	0.46	0.53
MIX_B	Image (to B from B)	7.67	7.20	7.67	N/A/	0.87	0.42	0.49	N/A

APPENDIX C.3 -- 1992 LAN/FASTPAC Testing Results -- By Configuration

		MEAN (secs)				STD. DEVIATION (secs)			
		NFSD = 0		NFSD = 8		NFSD = 0		NFSD = 8	
		LAN	FASTPAC	LAN	FASTPAC	LAN	FASTPAC	LAN	FASTPAC
Config.	Operation								
X-TERM (A)	Clean AAA_x	43.90				0.74			
X-TERM (A)	Clean AAAw_x		45.42			0.79	0.79		
X-TERM (A)	Clean AAB_x	67.40	87.25			0.70	0.62		
X-TERM (A)	Clean BAA_x		44.92				0.67		
X-TERM (A)	Clean BAB_x	68.00	87.00			0.47	0.74		
X-TERM (A)	Copy AAA_x	24.70				0.48			
X-TERM (A)	Copy AAAw_x		26.08				0.51		
X-TERM (A)	Copy AAB_x	43.30	57.64			0.48	2.06		
X-TERM (A)	Copy BAA_x		22.00				0.63		
X-TERM (A)	Copy BAB_x	50.00	73.17			0.00	0.39		
X-TERM (A)	Draw (to X from A)	28.90	28.00			0.74	0.74		
X-TERM (A)	Draw (to X from B via A)	29.60	30.33			0.52	0.78		
X-TERM (A)	Image (to X from A)	7.20	11.25			0.42	0.62		
X-TERM (A)	Image (to X from B via A)	17.80	32.92			0.63	0.29		
X-TERM (B)	Clean ABA_x		89.40				0.70		
X-TERM (B)	Clean ABB_x		47.40				0.84		
X-TERM (B)	Clean BBA_x		87.40				0.70		
X-TERM (B)	Clean BBB_x		47.30				0.67		
X-TERM (B)	Copy ABA_x		73.00				0.00		
X-TERM (B)	Copy ABB_x		21.90				0.57		
X-TERM (B)	Copy BBA_x		57.10				0.32		
X-TERM (B)	Copy BBB_x		25.90				0.32		
X-TERM (B)	Draw (to X from A via B)		29.60				0.52		
X-TERM (B)	Draw (to X from B)		27.00				0.00		
X-TERM (B)	Image (to X from A via B)		30.80				0.42		
X-TERM (B)	Image (to X from B)		7.00				0.00		

INFLUENCE OF ADDITIONAL NETWORK LOADING ACROSS LAN								
	CAA_t2	**CAA_t3**	**CAA_t4**	**CAA_t5**	**CAA_t6**	**CAA_t8**	**CAA_t9**	**CAA_ta**
clean AAB_x	67	70	70	75	75	91	90	91
copy AAB_x	44	45	45	45	45	47	46	46
draw (to X from A)	29	32	31	33	33	47	44	44
image (to X from A)	7	7	7	9	9	10	10	11
CAA_T2 = No Background Noise								
CAA_t3 = 10% Noise -- 1 each 1500, 60 and 256-byte packet, 1 frame per burst								
CAA_t4 = 10% Noise -- 1 each 1500, 60 and 256-byte packet, 3 frames per burst								
CAA_t5 = 20% Noise -- 1 each 1500, 60 and 256-byte packet, 3 frames per burst								
CAA_t6 = 20% Noise -- 3 1500-byte packets, 3 frames per burst								
CAA_t7 =								
CAA_t8 = 20% Noise -- 3 60-byte packets, 3 frames per burst								
CAA_t9 = 20% Noise -- 5 60-byte packets, 5 frames per burst								
CAA_ta = 20% Noise -- 10 60-byte packets, 10 frames per burst								
INFLUENCE OF ADDITIONAL NETWORK LOADING ACROSS FASTPAC								
Mean Times								
	traff_00	**traff_01**	**traff_02**	**traff_03**	**traff_11**	**traff_21**	**traff_22**	**traff_23**
Copy BBA	58.00	58.40	59.20	61.20	62.60	60.80	62.80	65.60
Copy ABB	22.20	24.80	26.20	28.20	26.80	24.40	26.60	30.20
Copy ABA	73.00	77.50	76.60	85.60	83.00	77.80	81.40	87.80
Image (from A to B)	31.20	35.20	37.20	40.80	38.20	35.00	39.00	43.00
Standard Deviations								
	traff_00	**traff_01**	**traff_02**	**traff_03**	**traff_11**	**traff_21**	**traff_22**	**traff_23**
Copy BBA	0.00	0.52	2.28	1.79	0.55	0.45	0.84	0.55
Copy ABB	0.45	0.63	3.42	2.59	0.45	0.55	0.55	0.84
Copy ABA	0.00	0.53	4.51	0.89	0.00	0.45	0.55	0.84
Image (from A to B)	0.45	0.42	5.26	5.02	0.45	0.71	0.00	1.00
traff_00 = No Background Noise								
traff_01 = 10% Noise, 3 1514-byte packets, 3 frames per burst								
traff_02 = 20% Noise, 3 1514-byte packets, 3 frames per burst								
traff_03 = 30% Noise, 3 1514-byte packets, 3 frames per burst								
traff_11 = 5% Noise, 3 60-byte packets, 3 frames per burst								
Note: Test with 10% Noise, 3 60-byte packets, 3 frames per burst --- Severe Delays -- Test abandoned								
traff_21 = 10% Noise, 1- 60,1- 256 and 2-1514 byte packets, 3 frames per burst								
traff_22 = 20% Noise, 1- 60,1- 256 and 2-1514 byte packets, 3 frames per burst								
traff_23 = 30% Noise, 1- 60,1- 256 and 2-1514 byte packets, 3 frames per burst								

INFLUENCE OF ADDITIONAL SERVER/MEMORY LOADING								
	BAA_t1	**BAA_t3**		**BAA_t2**				
Clean ABA	69.40	87.40	26%	131.50	89%			
Clean BBA	66.90	85.60	28%	125.00	87%			
Copy ABA	49.10	68.60	40%	120.00	144%			
Copy ABB	13.50	16.95	26%	24.00	78%			
Copy BBA	39.90	55.55	39%	95.50	139%			
Draw (to B from A)	29.30	30.30	3%	32.00	9%			
Image (to B from A)	19.00	24.25	28%	44.00	132%			
BAA_t1 = No Additional Server Loading								
BAA_t2 = Heavy Server/Memory Load (Constant Copying of large file running in the background)								
BAA_t3 = Medium-Heavy Server/Memory Load (Script of "Copy - Pause - Draw - Pause" Operations) running in the background.								

APPENDIX D

Effects of Varying the Physical Location of the GIS Software

Appendix D

EFFECTS OF VARYING THE PHYSICAL LOCATION OF THE GIS SOFTWARE

In cases where the same operations were performed on comparable usage configurations (i.e., AAA vs. BBB; BAA vs. ABB; ABA vs. BAB; and AAB vs. BBA), the *t-ratio* statistic was used to compare the respective mean values and determine if any significant difference existed between the corresponding samples. This statistic was calculated using the following formula:

$$t = \frac{\tilde{X}_1 - \tilde{X}_2}{\sigma_{\tilde{x}-\tilde{x}}}, \text{ where } \sigma_{\tilde{x}-\tilde{x}} = \sqrt{\sigma^2_{\tilde{x}_1} + \sigma^2_{\tilde{x}_2}} \text{ and } \sigma_{\tilde{x}_n} = \frac{\sigma_n}{\sqrt{N_n - 1}}$$

where:

$\tilde{X}_1 - \tilde{X}_2$ = difference between the means of the two samples;

$\sigma_{\tilde{X}-\tilde{X}}$ = standard error of the mean difference;

$\sigma_{\tilde{x}_n}$ = standard error of the mean of Sample n, where n = 1, 2;

σ_n = standard deviation of Sample n ; and

N_n = number of observations in Sample n.

A comparison of corresponding usage configurations indicates that — at least with the levels of memory possessed by workstations employed here — GIS performance generally remained similar *regardless of whether the application was stored locally or resided on the remote server across the FASTPAC cloud.* Differences in some corresponding mean values *were* interpreted to be statistically significant at the ".01" confidence level. Practically speaking, however, none of the corresponding times tested differed by more than 20% of the overall time involved and most differed by less than 8%.

APPENDIX D.1 -- Comparison of LAN/FASTPAC Results for Similar Configurations -- By Operation

		MEAN (secs)				STD. DEVIATION (secs)			
		NFSD = 0		NFSD = 8		NFSD = 0		NFSD = 8	
		LAN	FASTPAC	LAN	FASTPAC	LAN	FASTPAC	LAN	FASTPAC
Config.	**Operation**								
AAA	Clean AAA	44.30	45.70	45.73	44.50	0.67	1.06	0.88	0.55
MIX_B	Clean BBB	45.00	46.50	46.40	N/A	1.25	0.97	1.68	N/A
AAA	Copy AAA	24.70	26.00	32.00	32.67	0.67	0.00	0.38	0.52
MIX_B	Copy BBB	25.30	26.20	26.33	N/A	0.48	0.42	0.62	N/A
AAA	Draw (to A from A)	28.50	27.30	27.00	27.00	0.48	0.48	0.53	0.63
MIX_B	Draw (to B from B)	26.40	26.80	28.67	N/A	0.52	0.42	0.62	N/A
AAA	Image (to A from A)	7.60	8.00	7.93	7.50	0.52	0.82	0.70	0.84
MIX_B	Image (to B from B)	7.67	7.20	7.67	N/A/	0.87	0.42	0.49	N/A
MIX_A	Clean AAB	67.20	87.30	62.87	74.67	0.63	0.67	1.06	0.52
MIX_B	Clean BBA	66.80	87.40	68.40	84.90	1.03	1.43	0.74	0.32
MIX_A	Clean BAA	42.73	45.50	44.40	45.17	0.70	0.85	0.63	0.75
MIX_B	Clean ABB	44.33	46.60	44.47	45.00	0.50	0.70	0.64	0.82
MIX_A	Clean BAB	70.20	87.20	62.33	76.60	1.01	0.79	0.90	0.55
MIX_B	Clean ABA	68.40	89.00	66.00	81.80	1.17	0.67	0.93	1.87
MIX_A	Copy AAB	42.70	56.00	42.13	53.50	0.48	0.00	0.35	0.55
MIX_B	Copy BBA	39.90	57.10	37.93	47.80	0.32	0.32	0.26	0.42
MIX_A	Copy BAA	13.20	22.00	12.60	14.50	0.68	0.67	0.51	0.55
MIX_B	Copy ABB	13.50	21.90	12.93	15.20	0.53	0.57	0.59	0.63
MIX_A	Copy BAB	52.13	72.80	55.53	67.00	0.35	0.42	0.52	0.00
MIX_B	Copy ABA	48.10	73.10	51.07	61.40	0.32	0.32	0.46	0.52
MIX_A	Draw (to A from B)	29.47	29.90	27.47	28.00	0.52	0.32	0.52	0.00
MIX_B	Draw (to B from A)	29.30	29.50	28.87	29.60	0.67	0.71	0.64	0.52
MIX_A	Image (to A from B)	18.60	31.80	13.33	18.50	0.51	0.42	0.49	0.55
MIX_B	Image (to B from A)	19.00	31.30	13.27	18.50	0.00	0.48	0.46	0.53

T-Test		NFSD = 0		NFSD = 8	
		LAN	**FASTPAC**	**LAN**	**FASTPAC**
Clean AAA	**Mean #1**	44.30	45.70	45.73	
Clean BBB	**Mean #2**	45.00	46.50	46.40	
	Std. Dev #1	0.67	1.06	0.88	
	Std. Dev #2	1.25	0.97	1.68	
	Count N1	10.00	10.00	10.00	10.00
	Count N2	10.00	10.00	10.00	10.00
	M1-M2	-0.70	-0.80	-0.67	0.00
	Sig. x1	0.22	0.35	0.29	0.00
	Sig. x2	0.42	0.32	0.56	0.00
	Sig x1-x2	0.80	0.82	0.92	0.00
	T-Ratio	**-0.88**	**-0.97**	**-0.72**	**#NUM!**
	df	18.00	18.00	18.00	18.00
T-Test		**NFSD = 0**		**NFSD = 8**	
		LAN	**FASTPAC**	**LAN**	**FASTPAC**
Draw AAA	**Mean #1**	28.70	27.30	27.00	
Draw BBB	**Mean #2**	26.40	26.80	28.67	
	Std. Dev #1	0.48	0.48	0.53	
	Std. Dev #2	0.52	0.42	0.62	
	Count N1	10.00	10.00	10.00	10.00
	Count N2	10.00	10.00	10.00	10.00
	M1-M2	2.30	0.50	-1.67	0.00
	Sig. x1	0.16	0.16	0.18	0.00
	Sig. x2	0.17	0.14	0.21	0.00
	Sig x1-x2	0.58	0.55	0.62	0.00
	T-Ratio	**3.98**	**0.91**	**-2.69**	**#NUM!**
	df	18.00	18.00	18.00	18.00
T-Test		**NFSD = 0**		**NFSD = 8**	
		LAN	**FASTPAC**	**LAN**	**FASTPAC**
Clean AAB	**Mean #1**	67.20	87.30	62.87	74.67
Clean BBA	**Mean #2**	66.80	87.40	68.40	84.90
	Std. Dev #1	0.63	0.67	1.06	0.52
	Std. Dev #2	1.03	1.43	0.74	0.32
	Count N1	10.00	10.00	15.00	5.00
	Count N2	10.00	10.00	15.00	5.00
	M1-M2	0.40	-0.10	-5.53	-10.23
	Sig. x1	0.21	0.22	0.28	0.26
	Sig. x2	0.34	0.48	0.20	0.16
	Sig x1-x2	0.74	0.84	0.69	0.65
	T-Ratio	**0.54**	**-0.12**	**-7.98**	**-15.86**
	df	18.00	18.00	28.00	8.00

APPENDIX D.2 -- T-Test Comparison -- Times for Corresponding Operations

T-Test		NFSD = 0		NFSD = 8	
		LAN	FASTPAC	LAN	FASTPAC
Clean BAB	Mean #1	70.20	87.20	62.33	76.60
Clean ABA	Mean #2	68.40	89.00	66.00	81.80
	Std. Dev #1	1.01	0.79	0.90	0.55
	Std. Dev #2	1.17	0.67	0.93	1.87
	Count N1	10.00	10.00	10.00	10.00
	Count N2	10.00	10.00	10.00	10.00
	M1-M2	1.80	-1.80	-3.67	-5.20
	Sig. x1	0.34	0.26	0.30	0.18
	Sig. x2	0.39	0.22	0.31	0.62
	Sig x1-x2	0.85	0.70	0.78	0.90
	T-Ratio	**2.11**	**-2.58**	**-4.70**	**-5.79**
	df	18.00	18.00	18.00	18.00
T-Test		NFSD = 0		NFSD = 8	
		LAN	FASTPAC	LAN	FASTPAC
Copy BAA	Mean #1	13.20	22.00	12.60	14.50
Copy ABB	Mean #2	13.50	21.90	12.93	15.20
	Std. Dev #1	0.68	0.67	0.51	0.55
	Std. Dev #2	0.53	0.57	0.59	0.63
	Count N1	10.00	10.00	10.00	10.00
	Count N2	10.00	10.00	10.00	10.00
	M1-M2	-0.30	0.10	-0.33	-0.70
	Sig. x1	0.23	0.22	0.17	0.18
	Sig. x2	0.18	0.19	0.20	0.21
	Sig x1-x2	0.64	0.64	0.61	0.63
	T-Ratio	**-0.47**	**0.16**	**-0.55**	**-1.12**
	df	18.00	18.00	18.00	18.00
T-Test		NFSD = 0		NFSD = 8	
		LAN	FASTPAC	LAN	FASTPAC
Draw (to A from B)	Mean #1	29.47	29.90	27.47	28.00
Draw (to B from A)	Mean #2	29.30	29.50	28.87	29.60
	Std. Dev #1	0.52	0.32	0.52	0.00
	Std. Dev #2	0.67	0.71	0.64	0.52
	Count N1	10.00	10.00	10.00	10.00
	Count N2	10.00	10.00	10.00	10.00
	M1-M2	0.17	0.40	-1.40	-1.60
	Sig. x1	0.17	0.11	0.17	0.00
	Sig. x2	0.22	0.24	0.21	0.17
	Sig x1-x2	0.63	0.58	0.62	0.41
	T-Ratio	**0.27**	**0.68**	**-2.25**	**-3.86**
	df	18.00	18.00	18.00	18.00

APPENDIX D.2 -- T-Test Comparison -- Times for Corresponding Operations

T-Test		NFSD = 0		NFSD = 8	
		LAN	FASTPAC	LAN	FASTPAC
Clean AAB	Mean #1	67.20	87.30	62.87	74.67
Clean BBA	Mean #2	66.80	87.40	68.40	84.90
	Std. Dev #1	0.63	0.67	1.06	0.52
	Std. Dev #2	1.03	1.43	0.74	0.32
	Count N1	10.00	10.00	15.00	5.00
	Count N2	10.00	10.00	15.00	5.00
	M1-M2	0.40	-0.10	-5.53	-10.23
	Sig. x1	0.21	0.22	0.28	0.26
	Sig. x2	0.34	0.48	0.20	0.16
	Sig x1-x2	0.74	0.84	0.69	0.65
	T-Ratio	**0.54**	**-0.12**	**-7.98**	**-15.86**
	df	18.00	18.00	28.00	8.00

APPENDIX D.2 -- T-Test Comparison -- Times for Corresponding Operations

		NFSD = 0		NFSD = 8	
		LAN	**FASTPAC**	**LAN**	**FASTPAC**
Copy AAA	Mean #1	45.50	45.70	32.00	
Copy BBB	Mean #2	45.00	46.50	26.33	
	Std. Dev #1	0.71	1.06	0.38	
	Std. Dev #2	1.25	0.97	0.62	
	Count N1	10.00	10.00	10.00	10.00
	Count N2	10.00	10.00	10.00	10.00
	M1-M2	0.50	-0.80	5.67	0.00
	Sig. x1	0.24	0.35	0.13	0.00
	Sig. x2	0.42	0.32	0.21	0.00
	Sig x1-x2	0.81	0.82	0.58	0.00
	T-Ratio	**0.62**	**-0.97**	**9.84**	**#NUM!**
	df	18.00	18.00	18.00	18.00

T-Test		NFSD = 0		NFSD = 8	
		LAN	**FASTPAC**	**LAN**	**FASTPAC**
Image AAA	Mean #1	7.60	8.00	7.93	
Image BBB	Mean #2	7.67	7.20	7.67	
	Std. Dev #1	0.52	0.82	0.70	
	Std. Dev #2	0.87	0.42	0.49	
	Count N1	10.00	10.00	10.00	10.00
	Count N2	10.00	10.00	10.00	10.00
	M1-M2	-0.07	0.80	0.27	0.00
	Sig. x1	0.17	0.27	0.23	0.00
	Sig. x2	0.29	0.14	0.16	0.00
	Sig x1-x2	0.68	0.64	0.63	0.00
	T-Ratio	**-0.10**	**1.25**	**0.42**	**#NUM!**
	df	18.00	18.00	18.00	18.00

		NFSD = 0		NFSD = 8	
		LAN	**FASTPAC**	**LAN**	**FASTPAC**
Clean BAA	Mean #1	42.73	45.50	44.40	45.17
Clean ABB	Mean #2	44.33	46.60	44.47	45.00
	Std. Dev #1	0.70	0.85	0.63	0.75
	Std. Dev #2	0.50	0.70	0.64	0.82
	Count N1	10.00	10.00	10.00	10.00
	Count N2	10.00	10.00	10.00	10.00
	M1-M2	-1.60	-1.10	-0.07	0.17
	Sig. x1	0.23	0.28	0.21	0.25
	Sig. x2	0.17	0.23	0.21	0.27
	Sig x1-x2	0.63	0.72	0.65	0.72
	T-Ratio	**-2.53**	**-1.53**	**-0.10**	**0.23**
	df	18.00	18.00	18.00	18.00

T-Test		NFSD = 0		NFSD = 8	
		LAN	FASTPAC	LAN	FASTPAC
Copy AAB	Mean #1	42.70	56.00	42.13	53.50
Copy BBA	Mean #2	39.90	57.10	37.93	47.80
	Std. Dev #1	0.48	0.00	0.35	0.55
	Std. Dev #2	0.32	0.32	0.26	0.42
	Count N1	10.00	10.00	10.00	10.00
	Count N2	10.00	10.00	10.00	10.00
	M1-M2	2.80	-1.10	4.20	5.70
	Sig. x1	0.16	0.00	0.12	0.18
	Sig. x2	0.11	0.11	0.09	0.14
	Sig x1-x2	0.52	0.32	0.45	0.57
	T-Ratio	5.42	-3.39	9.31	10.03
	df	18.00	18.00	18.00	18.00
		NFSD = 0		NFSD = 8	
		LAN	FASTPAC	LAN	FASTPAC
Copy BAB	Mean #1	52.13	72.80	55.53	67.00
Copy ABA	Mean #2	48.10	73.10	51.07	61.40
	Std. Dev #1	0.35	0.42	0.52	0.00
	Std. Dev #2	0.32	0.32	0.46	0.52
	Count N1	10.00	10.00	10.00	10.00
	Count N2	10.00	10.00	10.00	10.00
	M1-M2	4.03	-0.30	4.47	5.60
	Sig. x1	0.12	0.14	0.17	0.00
	Sig. x2	0.11	0.11	0.15	0.17
	Sig x1-x2	0.47	0.50	0.57	0.41
	T-Ratio	8.53	-0.60	7.84	13.50
	df	18.00	18.00	18.00	18.00
T-Test		NFSD = 0		NFSD = 8	
		LAN	FASTPAC	LAN	FASTPAC
Image (to A from B)	Mean #1	18.60	31.80	13.33	18.50
Image (to B from A)	Mean #2	19.00	31.30	13.27	18.50
	Std. Dev #1	0.51	0.42	0.49	0.55
	Std. Dev #2	0.00	0.48	0.46	0.53
	Count N1	10.00	10.00	10.00	10.00
	Count N2	10.00	10.00	10.00	10.00
	M1-M2	-0.40	0.50	0.07	0.00
	Sig. x1	0.17	0.14	0.16	0.18
	Sig. x2	0.00	0.16	0.15	0.18
	Sig x1-x2	0.41	0.55	0.56	0.60
	T-Ratio	-0.97	0.91	0.12	0.00
	df	18.00	18.00	18.00	18.00

APPENDIX D.2 -- T-Test Comparison -- Times for Corresponding Operations

		NFSD = 0		NFSD = 8	
		LAN	FASTPAC	LAN	FASTPAC
Clean BAA	**Mean #1**	42.73	45.50	44.40	45.17
Clean ABB	**Mean #2**	44.33	46.60	44.47	45.00
	Std. Dev #1	0.70	0.85	0.63	0.75
	Std. Dev #2	0.50	0.70	0.64	0.82
	Count N1	10.00	10.00	10.00	10.00
	Count N2	10.00	10.00	10.00	10.00
	M1-M2	-1.60	-1.10	-0.07	0.17
	Sig. x1	0.23	0.28	0.21	0.25
	Sig. x2	0.17	0.23	0.21	0.27
	Sig x1-x2	0.63	0.72	0.65	0.72
	T-Ratio	**-2.53**	**-1.53**	**-0.10**	**0.23**
	df	**18.00**	**18.00**	18.00	18.00

APPENDIX E

Determining the Cost of FASTPAC Services Based on GIS Usage Monitoring Results

APPENDIX E

DETERMINING THE COST OF FASTPAC SERVICES IN A GIS CONTEXT

One of the objectives of the 1992 FASTPAC/GIS research was to provide a stronger link between GIS performance over FASTPAC and the component costs involved. As part of this year's research, the FASTPAC tariffs in place as of September, 1992 were loaded into a spreadsheet cost model. After a brief description of the model itself, the following sections discuss the cost implications of four particular issues raised during this year's research, including the costs of idling traffic on an X-terminal, projected Inter-LAN traffic levels at Conservation and Environment, and operational tradeoffs involved when dealing with large image files.

E.1 Spreadsheet Cost Model of the FASTPAC Tariff

A preliminary Telecom Price Schedule for FASTPAC services and a copy of the spreadsheet output is included at the end of this Appendix. The model has been developed to echo the various cost components of the approved FASTPAC Tariff released by Telecom Australia in April, 1992. To calculate the fixed charges involved, the user enters the number of connections required by the potential customer under the appropriate heading. (See Figure E.1.)

Variable costs are calculated on the basis of both standard "Unicast" tariffs and "Volume Link" discounts. Since these alternative tariffs depend on the distance between sites and the volume of data sent, estimates of both these parameters must be interactively entered by the user in the appropriate spot. (See Figure 15.) Again, the tables which calculate the resulting charges are based on FASTPAC tariffs released in September, 1992.

	ORGANISATION:	Vic C&E			
	TOTAL NO. OF SITES				
		Number of Sites / Type of Service			
		Number	2M Bit/S	Number	10M Bit/S
1)	Installation Charges				
	— NTU & Line	0	$0	2	$18,000
	— Add'n. Interfaces Conn. Later	0	$0	0	$0
	— Duplicated Line				
	— Dual Node Line				
2)	Yrly. Access Rentals				
	Prescribed Areas		$0	2	$97,728
	Non-Prescribed Areas		$0	0	N/A
	Yrly. Interface Rentals				
	— IEEE 802.3	0	$0	2	$18,552
	— IEEE 802.5	0	$0	0	$0
	— 2 Mbps Interface (alone)	0	$0	0	N/A
	— 2 Mbps Interface (in comb.)	0	$0	0	N/A

Number of connections and interfaces required is entered interactively by the User.

Totals are updated automatically.

Figure E.1

Entry of Fixed-Cost Components into Spreadsheet Model

FASTPAC Usage Charges						
EST. AVGE. MONTHLY DATA TRANSFER ACROSS FASTPAC						
FROM EACH SITE (MB per month)						
		≤3 km	4 - 50 km	51-500	500-800km	> 800 km
	No. of LANs					
Vic Pde.	1		35000			
Kew	1		10500			

Est. volume of data to be transferred from each site per month is entered interactively by the User.

Volume Usage Charges	Link	Length	Volume/mo.	Vol. Monthly Charge
	Vic Pde - Kew	20	35,000	$1028
	Kew - Vic. Pde.	20	10,500	$850

Figure E.2

Entry of Variable-Cost Components into Spreadsheet Model

Corresponding estimates of projected total annual charges based on both standard "Unicast" tariffs and volume-link discounted tariffs are then summarised at the bottom of the page. (See Table E.1.)

Table E.1

Cost Summary Format of Spreadsheet Model

Estimated Annual FASTPAC Charges		
	Standard	***Volume***
Installation	$18,000.00	*$18,000.00*
Annual Rental	$116,280.00	*$116,280.00*
Est. Annual Usage	$73,464.00	*$22,536*
Total	$207,744.00	*$156,816*

E.2 "Housekeeping Traffic" on an X-terminal

From the network monitoring research discussed in Section 3.4, it was suggested that an average traffic total of 75 kbytes over 6 minutes extrapolates to approximately 750 kbytes per hour per terminal of baseload traffic. Given these figures over an average month (7 hours per day, 20 working days per month), this would result in an average baseload of <u>approximately 105 Mbytes per month per terminal</u>.

Using the straight Unicast tariff charges, this 105 Mbyte figure translates to a monthly "base cost" of from approximately $10 to $165 per month per terminal working in single-shift operations depending on the distances between terminal and server. While the characteristics may vary, leaving an X-terminal logged into a remote host overnight would roughly triple this amount under the standard Unicast tariff. However, these "idling costs" may not have a significant influence on the customer's overall costs if charges are calculated using the "volume-link" tariff structure. The significance of these figures would ultimately depend on the nature of equipment and number of users at the customer's site(s).

Note: It must be stressed that these costs are based on underlying "housekeeping traffic" rates extrapolated from a limited number of observations. At best, they will be relevant to only the particular machines present on the network. If the Client wishes to obtain meaningful results with the necessary level of confidence, we strongly recommend that rigorous and repeated tests be undertaken on a wider range of equipment.

E.3 Costs of Current Inter-LAN Traffic at Victoria C&E

Tables 3.4 and 3.5 indicated the approximate levels of inter-LAN traffic to and from the central server at Victoria Parade. These observations were then interpolated to produce "normalised" rates of inter-LAN traffic in kilobytes per hour. Using a range of values obtained from these previous Tables, Table E.2 applies the standard Unicast and Volume Link tariffs to provide a comparison of estimated data traffic costs under each option.

The test names shown in the first column of Table E.2 correspond to those indicated in Tables 3.4 and 3.5. As can be seen here, inter-network usage is still relatively light by FASTPAC tariff standards. Even if the heaviest levels of traffic observed in the most recent tests were maintained, the current usage levels never exceeded the entry-level category of "volume-link" charges.

Table E.2

Comparative Costs of Data Traffic
Across 20 km. FASTPAC Connection
Unicast vs. Volume Link
(Based on traffic rates obtained from network monitoring experiments)

Test	From Victoria Parade (KBytes/hr.)	To Victoria Parade (KBytes/hr.)	Total Mbytes/mo. (~160 hrs/mo)	Monthly Unicast Cost $$/mo.	Monthly Volume Cost $$/mo.
Node-01	3380	2310	910	$164	$850
Node-03	3990	3770	1242	$224	$850
Node-04	420	360	125	$23	$850
KNode-01	180	80	42	$8	$850
KNode-02	24560	1470	4165	$750	$850
KConn-02	3750	5050	1408	$254	$850
KConn-04	100	0	16	$3	$850
KConn-05	680	350	165	$30	$850
KConn-06	7760	8770	2645	$476	$850

The cost figures shown in Table E.2 represent the distance from Victoria Parade to Kew (under 20 km.). Given the above traffic levels, there are no cases where users would experience any savings as a result of volume discounts. In fact, based on the figures and assumptions shown above, the heaviest traffic levels observed would have to be increased by approximately 15% and sustained at those levels in order for the "Volume Link" rate structure to be attractive to such users.

The low-speed network limitations understandably influence current user attitudes and production routines within the Department of Conservation and Environment to some degree. For example, remote users think twice before arbitrarily re-displaying a centrally-stored GIS graphics file on their terminal screen during an editing or database review session. Similarly, back-ups of various workstations on the network are handled locally at each site as opposed to centrally. Finally, the slower speed of even the ISDN links limit the amount of remote directories that can efficiently be mounted and transparently employed using NFS.

Based on experiences elsewhere, it can be assumed that network traffic will increase as: (a) the number of overall users increases; (b) greater usage of the central server and processing resources is made from remote sites; and (c) equipment and operating routines are modified to take advantage of faster communication between sites. In cases where "work-group" computing applications are emerging, the increases will be greater still. Predicting the degree of any such increases will depend on the individual working conditions and unique operating tradeoffs perceived in each organisation. While outside the scope of this research, it may be the subject of other research currently underway.

E.4 Dealing with Large Image Files

Output from the LAN Analyzer during performance testing sessions measured traffic between Server and Client during remote login sessions and normal sessions using remote, NFS-mounted directories. In most cases, traffic levels during remote login sessions were relatively small since most of the work was carried out on the server and only light text traffic (i.e., prompts and responses) came back to the Client.

The example which best illustrates the relevant differences between the two modes of operation took place during viewing of the 6 Mbyte image file (as discussed in Section 4.3.3). When the user was operating in a normal session and invoked a command to display a remote file, NFS would: (1) find the file on the remote disk; (2) *draw it all across into memory on the local workstation*; and (c) display it on the screen. In this particular case, this would mean that 6 Mbytes of data would be drawn across the FASTPAC cloud. At current FASTPAC "Unicast" rates, this single transfer would cost anywhere from $0.54 to $9.36, depending on the distance between sites.

By comparison, when the user "logged in " to the remote server directly and invoked a display command, *only the bit-mapped image required to fill the screen* (i.e., 1024 x 1024 pixels by 1 byte/pixel = ~1.05 Mbytes) was transferred across the network. Depending on the distances involved, <u>this</u> transfer would only cost between $0.10 and $1.72. While these differences may be smoothed out by the "Volume Link" tariff option, they may still be significant in cases where the potential customer has many remote users employing FASTPAC primarily to <u>view</u> many different image or bit-mapped graphics files in the course of a routine working day. If only limited local processing or manipulation of these file is required, it may pay the customer to upgrade his terminal server and central processing facilities rather than pay to have whole files constantly transferred across the system. Each case would have to be viewed on its own merits.