COMPUTER NETWORK ANALYSIS AND OPTIMISATION

by

GAVIN PETER RAY

A thesis submitted to the University of Plymouth
in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

School of Electronic, Communication and Electrical Engineering
Faculty of Technology

In collaboration with
AT&T ISTEL Ltd., Redditch, England

August 1993

COMPUTER NETWORK ANALYSIS & OPTIMISATION

GAVIN PETER RAY

ABSTRACT

This thesis presents a study and analysis of the major influences on network cost and their related performance. New methods have been devised to find solutions to network optimisation problems particular to the AT&T ISTEL networks in Europe and these are presented together with examples of their successful commercial application. Network performance is seen by the user in terms of network availability and traffic delay times. The network performance is influenced by many parameters, the dominating influences typically being the number of users accessing the network, the type of traffic demands they place upon it and the particular network configuration itself. The number of possible network configurations available to a network designer is vast if the full range of currently available equipment is taken into account. The aim of this research has been to assist in the selection of most suitable network designs for optimum performance and cost.

This thesis looks at the current differing network technologies, their performance characteristics and the issues pertinent to any network design and optimisation procedures. A distinction is made between the network equipment providing user 'access' and that which constitutes the cross country, or 'core', data transport medium. This partitioning of the problem is exploited with the analysis concentrating on each section separately.

The access side of the AT&T ISTEL - UK network is used as a basis for an analysis of the general access network. The aim is to allow network providers to analyse the root cause of excessive delay problems and find where small adjustments to access configurations might lead to real performance improvements from a user point of view. A method is developed to allow statistical estimates of performance and quality of service for typical access network configurations. From this a general method for the optimisation of cost expenditure and performance improvement is proposed.

The optimisation of both circuit switched and packet switched computer networks is shown to be difficult and is normally tackled by the use of complex procedures on mainframe computers. The new work carried out in this study takes a fresh look at the basic properties of networks in order to develop a new heuristic method for the design and optimisation of circuit switched core networks on a personal computer platform.

A fully functional design system was developed that implements time division multiplexed core network design. The system uses both a new heuristic method for improving the quality of the designs and a new 'speed up' algorithm for reducing times to find feasible routes, thereby dramatically improving overall design times. The completed system has since been used extensively to assist in the design of commercial networks across Europe.

Index

List of Figures

Glossary of Terms

APS	Accunet Packet Service
ASDS-E	Accunet Spectrum of Digital Services (Europe)
EINOS	AT&T Enhanced Network Optimisation System [Agarwal 1989]
Desensitivity	Cost range within which two routes of differing distance are considered
Destination	The final node of a route = end of a directed link connection between nodes
Elliptic Bound	Elliptical region about source and destination node within which searches for routes takes place
E1	A European 2Mbps trunk circuit, 32 x 64kbps channels
Hop	A component link in a route
IDNX	Integrated Digital Network Exchange, T.D.M. equipment from N.E.T.
IFR	InterSpan Frame Relay
Kilostream	64kbps digital circuit available in UK from BT
Link	The physically shortest connection directly between nodes
Megastream	2Mbps digital circuit available in UK from BT
Multiplexer	The equipment that concentrates many data channels into one of a single high capacity or subsequently de-concentrates the single channel into the many smaller components
N.E.T.	Network Equipment Technology, major networking equipment supplier
Node	The point in a network where a multiplexer is placed to give users access to the core network.
Outage	The loss of a network circuit for the purposes of passing error free data.
Path	The circuit over which two nodes connect - possibly via many links
Route	The ordered list of nodes visited in travelling from a source node to a destination node via many links
Route Finder	The algorithm used for searching for best routes between selected nodes
SA	Service Availability, hours of network 'outage' per month.
Source	The starting node of a route = the start of a directed link connection between nodes in a route

T.D.M.	Time Division Multiplexing, network capacity resource sharing scheme
Tail circuit	The PTT supplied circuit from the customer premises to the local PTT exchange.
Topology	The specific connection of links between nodes to form a network
T1	1.544Mb/s digital circuit available in USA from AT&T

Glossary of Symbols

b	Best cost in Route Finder algorithm.
B	Maximum hops allowed in the backup route.
c	Cost.
C	Capacity.
d	Desensitivity Value.
D	Demand, of traffic between two nodes.
E_{max}	Elliptic bound, maximum number of nodes in addition to P or B between which routes are searched.
f	Figure of merit, cost of trunk, path, including penalty costs.
F	Frame length.
g^{MAX}	Maximum flow rerouted through a trunk following the failure of another
g^{MIN}	Minimum flow lost from a link when a neighbouring one should fail
G	Maximum capacity flow through a trunk.following the worst case failure
H	Frame header size (characters).
h	Frame plex header size (characters).
H_M	Maximum number of hops allowed in a route.
l	Length of a trunk.
J	A selected route.
K	Decision variable, 0 or 1.
k	Single trunk setup cost.
K_0	Topology-wide trunk setup cost.
L	Number of trunks in a given topology.
m	Mean number of hops in a route.
N	Number of nodes in a topology.
n	Step n of a process.
p	Percentage desensitivity factor
P	Maximum hops allowed in a primary route.
Q	Unit cost per path.
RP	Revenue Potential of a network topology.

R	Route vector, list of all nodes within a route.
S	Sample size.
v	Step in the optimisation process.
V	Number of steps in the optimisation process.
W	Sum of trunk setup costs.
x	The distance between two nodes
α	Primary channel traffic on a trunk.
β	Backup channel traffic on a trunk.
θ	Percentage probability of delay bettering the threshold of τ mS
λ	Arrival rate.
μ	Frame processing time.
σ	Message length (characters).
τ	Delay threshold (mS)
Ω	Total allocated traffic on a trunk.

ACKNOWLEDGEMENT

The work contained in this thesis constitutes the results of a three year investigation into the various aspects of telecommunication network performance and design. All work was undertaken within the School of Electronic, Communication and Electrical Engineering, University of Plymouth, Plymouth, England under the supervision of Peter Sanders (School of Electronic, Communication and Electrical Engineering) and Dr Colin Stockel (School of Computing). The project began on 26 September 1988.

All work has been carried out under the industrial supervision of John Allen, Network Analysis Manager, Network Services Division of AT&T ISTEL Ltd, CDC, Redditch, England.

During the course of this research software packages have been developed to allow network analysis and optimisation to be performed. These packages have found profitable application in real network management situations.

All computer based work, except where stated otherwise, has been carried out on an IBM PS/2 model 70 personal computer based on the 16MHz 80386DX processor.

The software was developed using Borland Turbo Pascal v5.5.

This thesis was prepared using the Wordperfect Corporation wordprocessor Wordperfect v5.1 and printed on a SPARC laser printer using the Times Roman typeface.

All Trademarks are fully acknowledged.

DECLARATION

The work presented in this thesis is solely that of the author.

SIGNED ....Gavin Ray..........

DATE ....20 July '93..........

CHAPTER 1

Introduction to Network Analysis and Design

1.0. Introduction

The rapid expansion of the information technology industry over the last 25 years has resulted in great volumes of data being available to those who seek it. Merely determining the whereabouts of information is of little use, it is necessary to gain access to that data and read it or obtain a personal copy. This need to obtain information over distance, both great and small, has lead to the rapid development of computer communication networks covering areas from the size of a small office to whole continents.

Many computer networks now extend throughout national or even international regions and involve considerable sums of money in terms of equipment investment, installation costs and operational maintenance. The large expenditure on such networks makes the saving of even small percentages of outlay highly desirable, with the proviso that an agreed quality of service must always be offered to network users.

Computer network technology is not cheap, by its very nature it may involve many users and the connection of numerous items of equipment and sites. The larger the network the greater its cost and the more important it is to minimise expenditure to maintain its economic viability. One of the largest controllable costs of a network is that of the data transmission links between nodes. Given a number of cities, towns or nodes requiring connection to a network there are a great many possible ways of connecting those nodes and consequentially a wide range of expenditure options.

The particular choice of links between nodes may be determined by a number of factors, the primary criteria for link selection is the full provision of capacity to carry all the traffic required by each node. The commercial nature of computer networks results in the need to meet the primary design criteria subject to minimum cost constraints. The primary design criteria requires the knowledge of node to node traffic and therefore an important facet of network design is the understanding of user traffic profiles. A traffic profile is a statistical model that can describe the volume and arrival rates of data inputs and outputs due to any one user.

There are a number of technologies suitable for carrying data on computer networks. The specific choice being determined by traffic volumes, performance (measured in traffic delay and throughput) and cost criteria. The work conducted in this thesis is largely devoted to design requirements specific to networks operated by AT&T ISTEL throughout Europe. These networks are Infotrac, the Accunet Packet Service (ADS), InterSpan Frame Relay (IFR) and the Accunet Spectrum of Digital Services - Europe (ASDS-E) each of which is a trademarked product definition of a generic network service. Infotrac is a low speed (300bps to 2400bps) asynchronous network primarily designed to support viewdata services such as mortgage, insurance and holiday quotations as seen in the highstreet brokerages. ADS is the AT&T X.25 packet service, based on Bolt Bereneck and Neuman (BBN) switches, that operates throughout Europe. InterSpan is the new global frame relay service , operated using Stratacom switches, that is now available in Europe. ASDS-E is the high speed (2Mbps) network, based on Network Equipment Technology (NET) switches, available in 15 European countries and acts as the transmission network for Infotrac, ADS, and IFR.

This thesis describes methods for designing large scale computer networks suitable for carrying traffic on an international basis and concentrates on the use of time division multiplexing technology (TDM) from NET, which is currently the most versatile and economic system for carrying data from a range of differing applications. A method for creating traffic profiles is formulated and then, using this as an input to the design method, a procedure for designing feasible networks subject to lowest cost constraints is developed.

1.1. Layout of the Thesis

Chapter 2 serves as an introduction to the field of network technology and describes the primary factors in network design. A distinction between access and core network technology is drawn and the typical access and core configurations are discussed. The AT&T ISTEL UK network is outlined in its four phases of growth since the mid 1980's in order that the need for design and analysis effort be seen clearly. The network technology employed has advanced at each stage and the need for each change is highlighted.

In chapter 3 the issue of access network design is considered in isolation. The effect of different traffic profiles, due to a range of network users, is more formally analyzed in order that unified traffic related network performance measures can be created. A number of complex traffic profiles from a range of users are processed to produce a single graphical representation of the likelihood of given access network delay figures. The effect of a link level transmission protocol is included to show how network performance is modified by the use of such protocols.

The analysis and modelling of core networks is investigated in chapter 4. Not only are physical network parameters considered, but also the human perception of network topology. There are a number of perceptions that can assist in the design of networks, such as those based on the eye's ability to visualise clusters of nodes and perform concentrator location without resort to mathematics. However, there are also limitations in the way a network topology can be represented in a diagram and hence there are difficulties in interpreting topology diagrams. Chapter 4 also examines the linear programming approaches to core network design and concludes by introducing threaded design systems which are based on a connected series of route selections.

The complexity of the large scale computer network design problem is stated in Chapter 5 as being best illustrated by the work of Johnson, Lenstra and Kan [1978] who showed that the problem is NP-hard, having a time to solution function that is non-polynomial in N, where N is the number of nodes in the problem. The problem therefore requires solution methods other than brute force attack and an analysis is pursued with the aim of developing a heuristic network design method.

The application of linear programming methods to the network design task is considered inappropriate for large scale circuit switched network design under the cost and processing power limitations imposed in this project. The alternative threaded search method is developed and described in chapter 6 with details of the complete method. A rigorous design method is then developed which attempts to avoid many of the problems found in other systems, one of which is that they suffer from cost functions which may become trapped in 'local minima', where a topology design cannot be reduced in cost without undergoing considerable perturbation. The larger the required perturbation the greater the analysis time and, in the limit, the problem becomes unsolvable.

The results of a number of core network design scenarios are shown in chapter 7. Simple design methods are shown to produce poor results, yet as the design method is refined the cost of designs satisfying the same criteria is reduced. Concluding remarks on the effectiveness of the new design method are to be found in chapter 8.

CHAPTER 2

Network Technology and Techniques

2.0. Introduction

This chapter reviews the ways in which it is possible to link computer systems via a common network and examines the important factors that influence network designs. The systems discussed are typical of those pertinent to value added network providers such as AT&T ISTEL, in contrast to those that are provided by national Post, Telephone and Telegraph (PTT) suppliers. There are many differing technologies available to perform the function of data communication, each having both cost and performance profiles that make them suitable to specific applications. The fundamental properties of computer communication networks and their implications to cost and performance are discussed.

2.1. The Architecture of Computer Networks

A computer network is a system of data communication paths between interconnected nodes that allows the transfer of data between users. The network user may be a person at a terminal or a personal computer, PC, or alternatively a mainframe running application software. In addition to providing access to the network, a node offers switching facilities, directing traffic from incoming trunks to one of a number of possible output trunks. The users have differing traffic requirements in terms of volume and time, some users may periodically access the network for short transfers of data and there are those who require lengthy access periods for large volume data transfer. The network architecture must account for the nature of the user traffic and provide communication paths that furnish performance levels within bounds acceptable to the user at a cost acceptable to the network provider.

A network may take on many differing forms, from a simple ring to a complex mesh arrangement of links between nodes, the specific topology being determined by a number of parameters such as internodal link cost, data transfer volumes, time delay requirements, physical limitations and the data transfer mechanism.

The most common broad distinctions made between network types are local area networks

(LANs), metropolitan area networks (MANs) and wide area networks (WANs). LANs are typically defined as being less than 1 kilometre across and are becoming increasingly common offering distributed computing, within campus, corporation and public sector sites, taking over from the traditional centralised mainframe based processing. MANs are described [Dettra 1991] as being 'metropolitan (city) area sized high-bandwidth dual ring networks that allow data from individual locations (normally LANs) to access the network, to be transported across it, and to egress at a distant destination'. WANs are usually networks of greater size than a MAN and may interconnect a number of cities, towns, LANs and MANs. The number of networks throughout Europe and America is increasing rapidly as commercial communications traffic grows and inter-LAN connections become increasingly desirable. The business advantage of inter-LAN connection is generally attributed to the greater possible speed of decision making and action, based on improved inter-personnel communications through the use of electronic mail and greater sharing of data within documents, spreadsheets and databases.

2.1.1. Local Area Networks

In the strictest sense a local area network may be any system of data communication within a single area owned by the network operator. This could cover mainframe based proprietary systems such as IBM's SNA and Digital Equipment Corporations's DECnet though these systems will also operate over wide area networks and are usually thought of as WAN technologies. The term LAN is most commonly confined to the shared media broadcast systems, Ethernet and IBM Token Ring [Schwartz 1987]. Token Ring is based on a 16Mbps ring bus and Ethernet is normally implemented using a 10Mbps broadcast bus. In terms of topology, the single resilience attribute of the bus network is that if a break should occur then only communications across the break are severed, communication on each remaining segment may continue. The Token Ring network will fail if any single link is broken. The bus architecture has the advantage of being simple to configure, low in cost and physically easy to extend by adding new lengths of cable and terminating connections. This makes it an attractive option for the connection of multiple personal computers, PCs, in many commercial environments.

A broadcast system suffers where a number of heavy users require transmission within a small area of the network, since this will congest the entire system. This problem becomes significant over large LANs where, if unrestricted, a small number of users may congest the network to the detriment of the majority of other users. For this reason bus networks are often segmented using equipment called bridges or routers between segments. Bridges and routers can identify whether traffic is destined for the local segment or whether it should be passed to an external segment. Bridges only operate between two segments whereas a router can perform the same function between many segments. This segmentation is usually performed to maximise intra-segment traffic and minimise that which is inter-segment. LANs are becoming increasingly important customers of wide area networks since the traffic between segments is usually lower in volume than that within segments and may therefore be carried on lower speed WAN connections, for example 64k inter-segments trunk speeds compared to 10Mbps intra-segment speed.

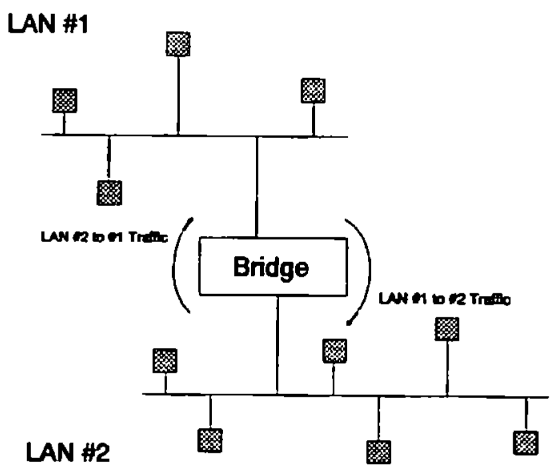


fig 2.1 Large scale bus with interconnected segments

2.1.2. Metropolitan Area Networks

Metropolitan area network technology is a new and expanding field which has been standardised by the IEEE as 802.6. A MAN covers an area larger than that of a single LAN, typically within the extent of a single city or town. MAN technology is based on high speed packet switching and fibre-optic media in a dual ring topology. An example of a MAN system is the Switched Multimegabit Data Service (SMDS) [Cox et al 1991]. This is effectively a 'bandwidth on demand' system offering high data rate in the same 'dial-up' manner as the voice telephone system. User lines are intended to interface directly to US T1 (1.544Mbps) and T3 (45Mbps) or European E1 (2Mbps) and E3 (34Mbps) links with provision for later use on 150Mbps when SONET (Synchronous Optical NETwork) [Johnson 1991] technology becomes commercially operational. MANs are a source of traffic demands for wide area networks.

2.1.3. Wide Area Networks

A wide area network is typically considered to cover a distance of more than one kilometre and may have connections to many forms of network technology including MANs and LANs. A wide area network is used to link users over great distance and will typically have a mesh topology between the major WAN nodes. Of prime importance to WAN providers is the provision of resilient network connections. In a large network, failures, though rare, are potentially damaging to customer business and mesh topologies offering multiple paths to all users are required to ensure reliable communications. Typically designs cater for single link or node failures, though more resilience is not uncommon but more expensive. The selection of links between nodes is based on minimising cost to carry the offered network load. Value added network providers, VANs, operate by offering communication capacity and management functions to the customer at a price the they could not themselves match. The economies of scale available to a VAN operating a large shared service are usually greater than individual customers can achieve. They also have the advantage of established links with international PTTs and the multi-lingual and multi-currency trading procedures. Long distance channels and expensive equipment may be shared by using one or a number of differing technologies to offer services to customers suited to their traffic requirements. Many large corporations have their own WANs connecting thousands of users over many sites, examples being ICI, IBM, Hewlett Packard, British Gas plc, Power Gen plc and one of the largest being General Electric Information Systems (GEIS).

WAN users are usually concentrated in a number of areas such as towns or geographically diverse commercial organisations. Dial-up telephone access may be offered by the WAN provider using banks of modems in city centres to give a wide potential for remote access to the network. Such modem access is typically provided in many towns and cities in order that the majority of network users need only make a local call to gain access. Wide area networks are usually required to offer a range of remote users access between each other and to centralised databases, messaging or information systems. The services offered by such systems are many and varied with new offerings such as electronic mail and voice messaging being high growth areas.

With the growth of LAN based networks it is also common for corporations located across a number of sites to require interconnection of their Ethernet and Token Ring LANs. This may be achieved by providing dedicated channels across the WAN to interconnect a number of bridges and or routers.

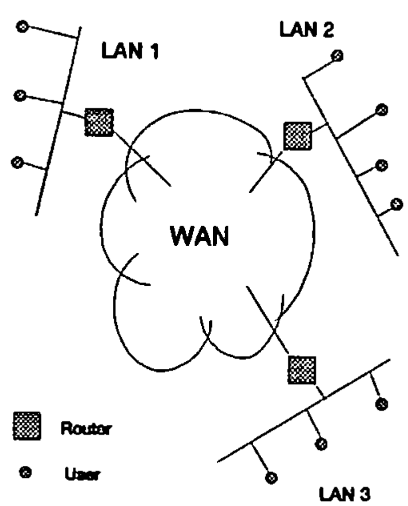


fig 2.2 LANs become access networks to WAN connections

Wide area networks are also provided by the PTTs themselves allowing public access to data networks for those organisations and individuals unable to justify the cost of establishing their own networks. These public digital networks also offer a range of services that users may subscribe to such as legal, news, medical, financial and technical databases. Examples of national high capacity X.25 systems are the UK based AT&T Accunet Packet Service, ADS [AT&T ADS Product Guide, 1992], the British Telecommunications packet switched network PSS (Packet SwitchStream) [Lane 1987], the American Government run Internet [Frank, Kahn, & Kleinrock 1972], US commercial Tymnet [Schwartz 1987], the French Transpac [Schwartz 1987] and the Canadian DataPac [Schwartz 1987] networks.

2.1.4. The Hierarchy of LANs, MANs and WANs

As networks grow large the problem of controlling the routing between vast numbers of nodes becomes significant. If a centralised routing system were to be employed then the sheer volume of routing data that would need to be transmitted from the central control computer to each router would consume large volumes of network resource. However, if each router were left to determine its own route to each and every other router then the flood of traffic required to establish contact would also consume an unacceptably large amount of network transmission capacity.

There is a square law increase in the processing and memory requirements placed upon routers for determining routing tables. With such complexity in large networks it is vital that some form of decomposition, or breaking down, of the topology is made to aid or even make possible both the operation and modelling of the network.

For these reasons a hierarchical network design is desirable which limits the bounds of any one routers 'world'. This has the effect of breaking computer communication networks up in the same regionalised way as found in voice communication networks.

The diagram below illustrates a simple hierarchical network. The important feature is that all routing between levels must be done at the highest common level. This simplifies routing at each level since any routing required to any node not on the same level is only required to the local 'gateway', the link to the level above. The diagram therefore represents the logical operation of routing between LANs at the lowest level, MANs at the intermediate level and the WAN at the top-most level.

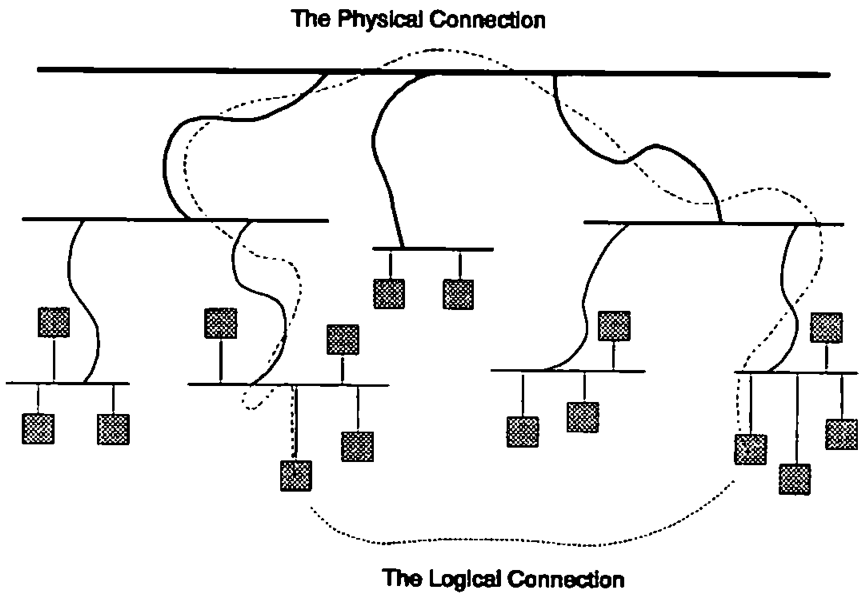


fig 2.3 A multi-level hierarchical network

Such hierarchical networks are becoming a practical reality as internetworking becomes increasingly common and the wide area coverage now extends around the world for many corporate networks. Large organisations are seeing the benefits of the synergy resulting from the joining together of their previously isolated departmental networks and subsidiaries. Gitlin & Kaufield [1988] state that companies are integrating their networks to gain advantages in the increasingly competitive global economy. They give figures stating that the explosive growth in installed LANS has lead to an average annual growth rate of 69% from 100,000 in 1985 to 800,000 in 1988 with a projection of 3.6 million by 1992. The linking together of local area networks is now made possible by the use of wide area connections between bridges and gateways, which make it possible to join X.25, Ethernet and Token Ring [Datacom, March 1991]. Perhaps the largest current hierarchical network is the American Internet [Schwartz, 1987] which now joins together many thousands of host computers, numerous local area networks and offers gateways to other major international networks such as GTE Telenet, Canadian Datapac and Tymnet.

2.2. Shared Data Systems

It would not be economically feasible, or indeed physically practical, to provide single dedicated communications paths between each and every pair of people requiring inter-communication. For this reason a range of methods have been devised to allow many users to share common transmission media. There are three predominant data switching systems in use today over shared communications resources. These are message switched, circuit switched and packet switched networks [Stallings 1991].

A shared infrastructure network offers connectivity between a large number of possibly disparate users by shared links between network nodes. In establishing communication between users a number of adjoining links may be chosen to construct a communication path between users and a selection method is required to define this path. The terms path and route tend to be used interchangeably in this field but the consensus appears to refer to routes and routing methods in general terms and communication paths for particular instances.

The predominant difference between the various network sharing systems is the type of traffic for which each is intended. Packet switching systems are implemented for traffic that is ideally characterised by random short bursts of data interspersed by periods of inactivity. Time division multiplexing (TDM) systems are designed to offer user selectable 'virtual channels' though a mesh network, with a fixed rate channel capacity.

2.2.1. Time Division Multiplexing

Time division multiplexing at its most basic level offers a means for sharing a single communications path between a number of users by apportioning transmission timeslots to each user. In early systems these timeslots were of fixed length and as such offered fixed channel capacities. Typically a 2Mbps channel could be divided into 32 channels of 64kbps each. The simplest mode of operation is referred to as synchronous TDM because each input channel is allocated equal length timeslots on the output channel and the receiving TDM operates in synchronisation with the transmitter. The timeslots are directed to each corresponding output path and transmitted at the lower rate.

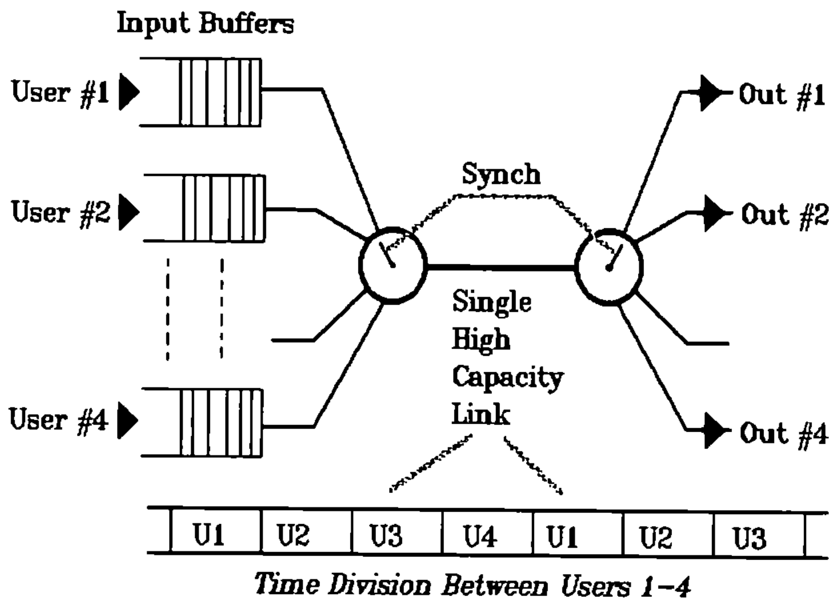


fig 2.4 Simple synchronous time division multiplexing

The synchronous TDM system is not efficient when any of the inputs are idle, the high capacity link will be required to transmit null data characters - in order to maintain synchronisation of each end of the channel. Should synchronisation be lost then data corruption will occur. A more efficient system uses a form of demand driven slot allocation where transmission time is allocated to data from the busiest inputs on a statistical basis, hence the term statistical time division multiplexer or stat-mux.

This makes improved use of the transmission link and allows the connection of more bursty traffic (high ratio of peak traffic levels relative to the mean level) links than would be possible with a fixed TDM system. This type of traffic is typical of interactive user sessions on terminals and PCs.

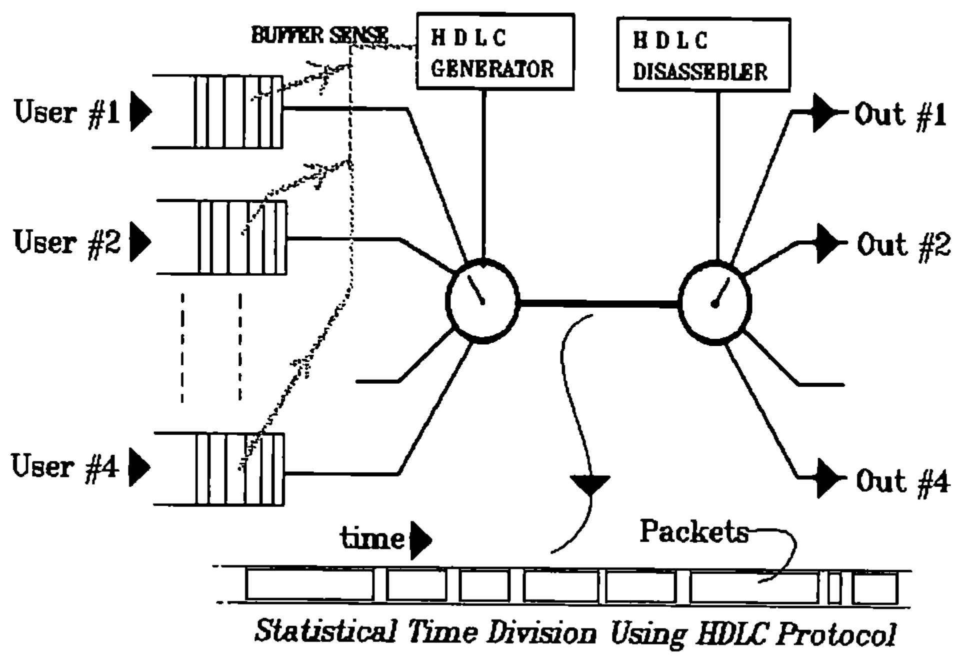


fig 2.5 The statistical multiplexer

Because the data is transmitted in variable length time frames the receiver requires a means for identifying from which input each time frame originated. This then determines to which output each frame is directed. This is achieved using a framing protocol which appends the data with input address and link management data. Typically the High-level Data Link Control protocol, HDLC, is used, [Stallings 1991]. This type of system does not guarantee capacity to each channel, being traffic dependent.

TDM systems originally only operated on a point to point basis but time division multiplexer technology has advanced further with the introduction of intelligent systems that offer end to end routing over mesh topologies and variable bandwidth allocation. The Network Equipment Technology Inc (NET, Redwood City, Ca, USA) IDNX (Integrated Digital Network Exchange) intelligent TDM system, as used by AT&T ISTEL, can allocate channels of any size (less than the trunk bandwidth) into trunks of capacity ranging between 64kbs and 34Mbs. NET describes the equipment as a transmission resource management (TRM) system because it offers dynamic route control in the event of link failure and can allocate transmission bandwidth on demand. There is no fixed time-slot size and so it is possible to carry efficiently a wide range of different input data rates.

In contrast to the HDLC system the NET IDNX guarantees a fixed capacity channel, using a small portion of each trunk capacity for internodal management purposes. User channels may be at any of the standard communication rates of 1.2kbps, 2.4kbps, 4.8kbps, 9.6kbps, 14.4kbps, 19.2kbps, 64kbps, 128kbps to 1024kbps and above.

An advantage of a TDM system is that once a virtual circuit is established communication operates at the assigned channel capacity with delays only attributable due to channel buffering and propagation times. Propagation delays are significant over international networks where 1mSec per 160km is expected [BT Engineering, Edgbaston, England]. The disadvantage of circuit switching is that the entire channel is allocated for the duration of the call, thus for bursty traffic (periods of heavy activity interspersed with periods of no activity) such as that found on most terminal-host calls the allocated channel capacity is wasted while no data is in transit.

A further advantage of the latest TDM systems is the range of features offered by the 'intelligent capacity management'. The systems are able to detect the failure of internodal trunks and, where possible re-establish each virtual circuit by alternate paths automatically. The selection of the channels to be carried when capacity is limited is determined by a system of call priorities, whereby calls are rerouted according to the assigned priorities, highest first.

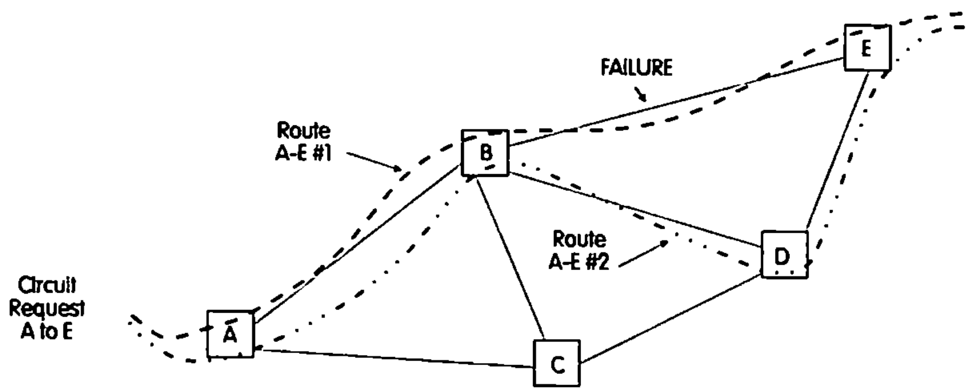


fig 2.6 Intelligent TDM establishes a new route after trunk failure

The economics of TDM systems are such that node costs are low in comparison to international circuit costs. For a value added network provider the costs of providing customer channels between 1.2kbps and 64kbps as TDM channels at fractions of an international 2Mbps circuit is considerably lower than a customer seeking the same capacity, dedicated channels from PTTs. In addition, the intelligent TDM systems are able to offer resilience to trunk failure by the automatic establishment of backup paths, this and the attendant network management and reporting functions are features for which customers are prepared to pay a premium. For these reasons intelligent TDM offers a viable means of supporting a wide variety of services if cost efficient TDM topologies can be developed.

2.2.2. Packet Switching

When traffic requiring a communication path between users is not continuous, but has a time variable load it is possible to implement a capacity sharing scheme that utilises the periods of inactivity between traffic bursts. The transmission of many non-continuous data streams across a single link may be achieved by using packet switching, first proposed in the early 1960's by the Rand Corporation [AT&T Training Course, X.25 and Packet Switching]. The most common implementation of the technique is defined by the CCITT international protocols collectively known as X.25. It allows the entire capacity of a link to be allocated to multiple users in a dynamic way, dependent upon the traffic demands of each user. Assuming the total mean user demand does not exceed the line capacity available packet switching offers an efficient way for multiple users to share a common transmission resource.

The data from a number of users is buffered and formed into packets by a processor implementing the packet switch protocol. The processor's tasks include the handling of all input/output buffering, connection monitoring, error checking and traffic flow control. The packets consist of discrete data streams containing packet sequence markers, addressing codes to indicate for which destination each packet is intended, the original user data and checksumming codes for error detection.

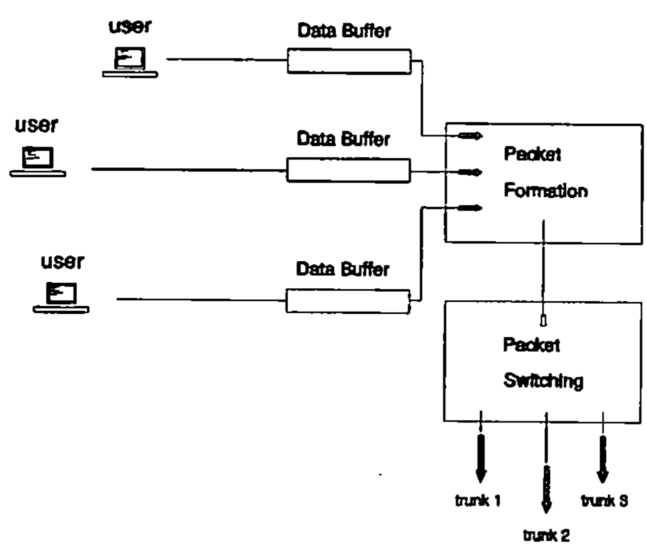
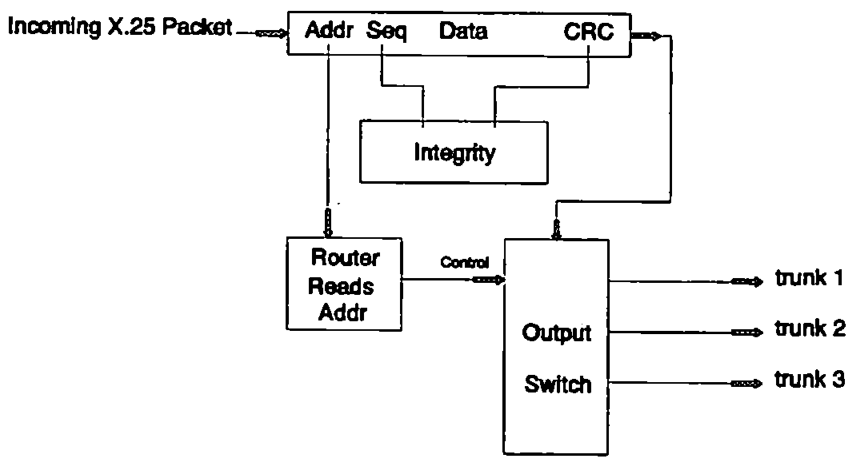


fig 2.7 Packet formation and switching

The basic packet formation arrangement represents only part of the packet switch system. It is also necessary to provide a routing system, whereby the nodes within a complex mesh network are able to direct packets along routes between source and destination nodes.



Addr = Address, Seq = Sequence Number

fig 2.8 The X.25 router reads address field to select output path

When a node has more than one outgoing link a selection must be made for each possible destination node. Packet systems usually decide upon the 'best' route by monitoring packet queuing times on each connected outward trunk and communicating to all other nodes the results at suitable intervals. A routing table may then be maintained either at each node or by a central network controlling node, holding a list of the minimum delay paths between each source and destination node.

2.3. Communication Link Capacity

Connections between nodes (cities and towns) are made by the use of communications links provided by local telecommunications suppliers (in UK such as BT or Mercury - or in the USA by AT&T, Sprint, Tymnet). Various capacities are available and the unit distance cost rises with increasing capacity, as described in 5.3.2. For national circuits (those originating and terminating in the same country) the cost is made up of an annual charge linearly related to the link distance plus a fixed circuit setup charge and further fixed annual rental. Typical line capacities are 300bps, 600bps, 1200bps, 2400bps, 9600bps, 19200bps for analogue circuits and 64kbps, 128kbps, 256kbps, 512 kbps, 1Mbps, 2Mbps, >2Mbs for digital circuit terminations.

The provision of international circuits is complicated by the differing states of technological investment and infrastructure around the world. Even by confining the scope of this investigation to European countries it is found that, while analogue circuits are readily available from 300bps to 19,200 bps, there are widely differing availabilities and costs of digital circuits. In designing a network to cover Europe for instance, the availability of the various circuit capacities must be taken into account. Countries such as Spain and Austria can only offer international circuits at 64kbps or 2Mbps whereas France, Luxembourg and Belgium can offer rates of 64kbps, 128kbps, 256kbps, 512kbps, 1Mbps and 2Mbps.

There are economies of scale involved in line rental for the various data communication link capacities available from PTTs around the world and this means that the cost per bit falls as line capacity increases. This is because the large PTTs and major communication capacity suppliers such as AT&T and MCI operate very high capacity satellite and fibre optic transmission systems between major commercial areas around the world. The higher the capacity the user requires, the less processing and capacity division equipment is required between the main arterial links and the user. This translates into lower overall costs due to the lower capital equipment and administration requirements.

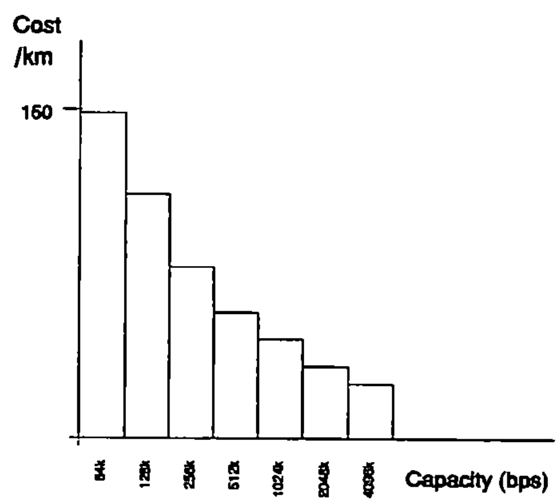


fig 2.9 Typical UK cost function of data links

In Europe the 2Mbps circuits are provided by the various PTTs in a number of different forms but generically referred to as E1 circuits. Out of the 2048kbps provided at the circuit termination the usable bandwidth is 1984kbps since one of the 32 64kbps timeslots is reserved for link synchronisation. In France a further 64kbps is reserved for PTT signalling and the usable capacity is therefore 1920kbps. Their equivalent in the US is the 1.5Mbps circuit referred to as a T1. Higher speeds are becoming available as 'standard offerings' (PTT standard services) at 34Mbps E3 in Europe and T3 at 45Mbps in the US.

By way of example, the following table illustrates the current prices of European links [Heywood, 1991], when comparing such prices the circuit costs are based on adding one connection to the local exchange to the international circuit cost. This is usually assumed to be a 2km local connection for approximation purposes.

Country	64kbps Local	64kbps 100km	2Mbps local	2Mbps 100km
Belgium	$389	$3,457	$3,566	$12,099
France	643	3,403	1,399	11,388
Germany	377	4,904	2,263	49,044
Italy	215	1,479	1,542	14,695
Luxembourg	123	948	615	4,141
Netherlands	278	1,282	1,393	5,854
UK (BT)	449	766	812	4,108
UK (Mercury)	200	545	700	3,415

Table 2.1 European link costs (1991) [Heywood, 1991]

The UK costs shown for Mercury are significantly cheaper than BT due to the predatory pricing policy resulting from the current duopoly and BT's dominant position.

Detailed tariff information is compiled by large consultancies such as Logica and Intelidata who produce Tarifica, which is a complete reference list of international circuit costs.

2.4. Two Tier Hierarchical Networks

Large scale commercial networks can often be reduced to two major sections for the purposes of analysis. The first section is that part of the network to which the user gains entry and the second is the internal network structure that carries the data over any significant distance from the point of access to the intended destination.

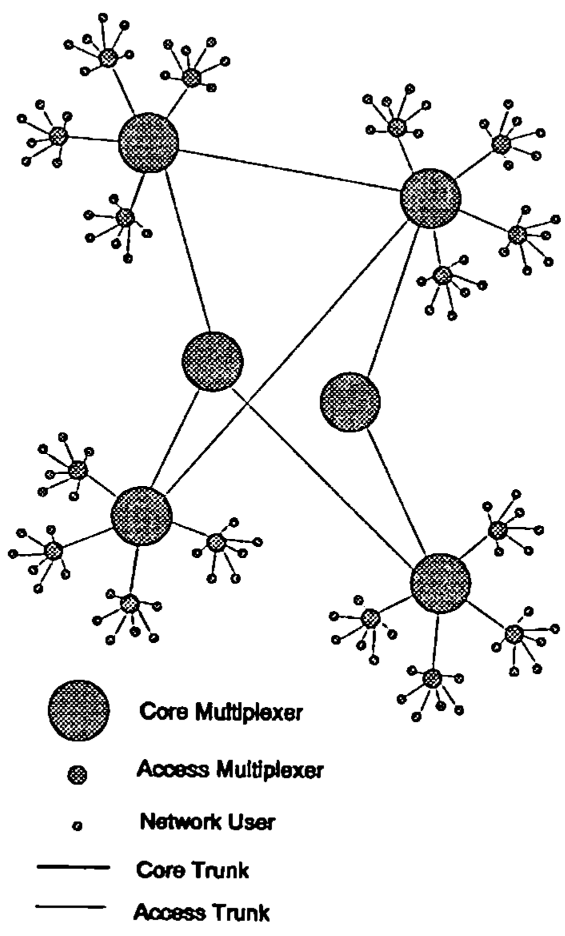


fig 2.10 A hierarchical network with distinct access and core sections

This distinction between access and core network is made because the very nature of the network structure lends itself to partitioning the problem of analysis. From the perspective of the end user, the core network is invisible, all that is apparent is the local connection and the far receiving end. The overall effect of the other network users is to present an aggregate load on the network which results in delays and throughput limitations. For a large network no one user should significantly impact another. For this reason, when analysing the access network the core may be modelled as a single entity with simplified parameters based on the aggregate load.

Similarly for analysis of the core network, the access network may be reduced from multiple inputs to a single input at each node where the total traffic is a statistical model representing the sum of the connected users.

Access to network resources is predominantly affected by three main factors:

the network capacity in terms of the number of access points and line data rates;

the number of calls (e.g. expressed as mean per hour);

the traffic volume of the calls (e.g. expressed as mean characters per second).

There are very few common computer applications, in terms of capacity, that require network services to provide permanent dedicated end to end capacity; batch file transfer and telemetry systems being the most likely such examples. Most network traffic is generated by computer users working either interactively with database query type systems or with the so-called productivity software packages. This term encompasses all programs such as spreadsheets, project management systems, personal organisers and computer aided design and analysis software. As a result, at the access layer of a network a packet based protocol is likely to be the most efficient. This allows a high capacity for short periods rather than a low constant capacity which may lead to throughput bottlenecks during short periods of high traffic demand.

The probability of a user gaining access to the network is affected by the number of calls per hour and the mean time a call lasts. With real networks there are a finite number of access points, referred to as ports. When the network is busy there may be many users attempting access and all the available ports may be busy. Any additional callers to the network will find that they cannot make an access because no ports are available and each call is therefore 'blocked'. This causes a total failure of the particular network access attempt. An important performance parameter of a network configuration is one of delay due to queuing. Where many traffic loads are coincident at a network access point they may be processed in turn at a rate slower than the arrival rate. Given that a call is successful in gaining access to the network the traffic submitted to and returned by the called party may travel through a number of switches and/or multiplexers. In joining the many other inputs to each multiplexer or switch, the traffic will be queued in buffers to await processing and transmission times will suffer a measure of delay while other prior arrivals are despatched.

2.4.1. Isolating Access Networks

The core is transparent to the access network and the access network is considered to be the entry and exit points for the user. The properties of an access network are topology, link capacities and access facility. The access facility may be a modem, a leased line or some form of local area network such as Ethernet or Token Ring. There may be queuing delays as data is funnelled in and out of the access multiplexers. Since the core network is transparent to the access layer yet may introduce its own queuing and switching delays, a specific fixed mean delay factor may be attributed to the core. This mean figure is determined from a separate analysis of the core in isolation. The advantage of this is that it is possible to carry out core and access network design and analysis in isolation. Each is a complex task in its own right, there is considerable reduction in effort if access performance is assumed fixed for core design and core performance is fixed for access design. The problem with such a partitioning method is that an artificial division cannot be imposed on a network that does not have a clear division between access and core functions. This situation is likely to occur in corporate networks that have grown sporadically over a period of time with no coherent strategy, this is not typical of strategic networks supported by value added resellers.

It is the topology of the access network which dictates the multiplexer to which each user is connected and the siting of multiplexers. It is common for the access layer to also be hierarchical, employing a number of low level multiplexers to groups of users within access regions. A smaller number of higher level multiplexers then concentrate the regional access to provide the interface to the core network.

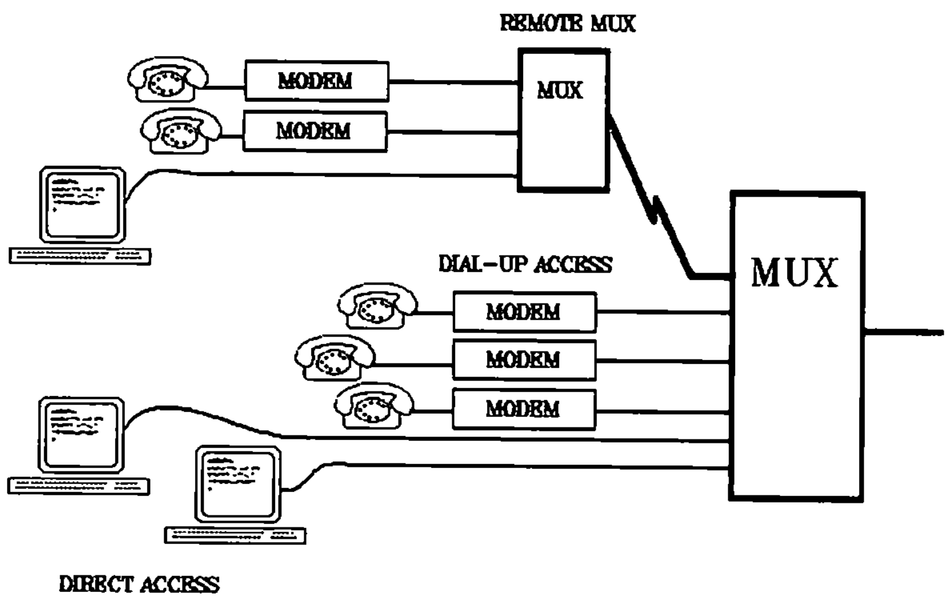


fig 2.11 A typical access configuration

2.4.2. Isolating the Core Network

The access network may be modelled as single point data flows into and out of the core network. This then provides simplified traffic inputs to a core network analysis. In a circuit switched core network there are no queuing delays, but 'blocking' can occur when more channels are required than are available. Blocking may occur at any point within the core circuit switched network and if alternate paths are not available, then the call fails. In a packet switch core network there can be no blocking (unless logical channels are exhausted) but queuing delays will become significant if excessive load accumulates.

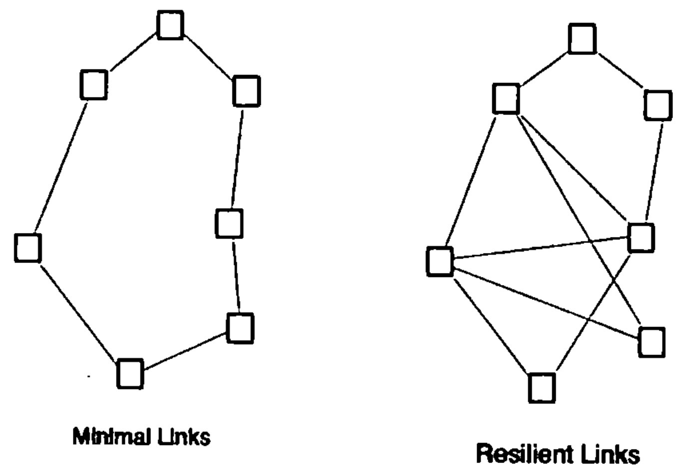


fig 2.12 Examples of possible core network configurations

2.5. Network Traffic

Traffic is the general term applied to the data that traverses a computer network. The data represents the transfer of information between users and the supervisory information sent between network control systems. The traffic has attributes of instantaneous volume and period cycles of zero demand. The requirement to transmit data throughout the network may arise at any time and the volume of data may take on any value according to the particular user. The volume of traffic due to one user may be of any size. At the smallest extreme, a user might be asked to type in a single key selection, to choose from a menu of services; at the other, they might issue a request across the network to a remote computer to transmit the entire contents of the Encyclopedia Britannica. Irrespective of the size of traffic demand, it will pass through the network at a rate limited by the network capacity and traffic due to other users at each intervening point.

2.5.1. Demands Across Networks

In general, networks have client parties who issue requests for information and server parties that return the required data. A user requesting data from a database is a client and the database manager is the server, whereas when a user feeds new data the roles are reversed, the database manager becoming the client. The demand for network services can generally be classed as being for interactive or file transfer operation. For the AT&T ISTEL network the interactive service is typified by remote access computer systems such as those that provide insurance or mortgage quotations to high street brokers and building societies. The users have small videotext terminals and use modems or direct links to access distant host computers, typically large IBM or ICL mainframes running large database systems, via the network. The advantage to the user is that a telephone call or a direct connection is only required to the local network node, while the host computer may be many miles away connected to a remote network node. This results in operational costs being kept to a minimum. When a user connects to the local network a call is established to the remote computer host and a session is said to have started. A session consists of a series of queries in the form of keypresses being sent by the user to the host and a number of streams of answering text returned from the host. Thus a session consists of a few bursts of characters from user to host each followed by a usually longer series of streams of text back from the host forming the reply to the query. Typically sessions of this type take between two and ten minutes as the customer issues a range of queries depending upon the type of data required. Typical data rates used between user and local network node are 1200bps, 2400bps and sometimes 9600bps.

The alternative view of interactive sessions is that supported by the Ethernet and Token Ring type LAN systems providing remote software applications. A single copy of a spreadsheet or accounting package may be held on a file server and remote users may access the network for copies of the software. The software is supplied in portions across the network as required, this is usually in concert with sophisticated operating systems like Unix and PC based LAN Managers.

The second major class of network service, the file transfer; is less common for large national network providers but serves a growing market and is becoming more significant. As the name implies this involves the transfer of large files in a steady stream, the response time of the network to file transfer being usually unimportant, while the throughput is critical. The data rate between user and network is preferably as high as possible and ranges from 64kbps to over 1Mbps but this is not provided through a low rate multiplexer such as for interactive users, but often via a direct connection to one of the core level multiplexer inputs in the case of AT&T ISTEL.

The arrival of user requests at a network have certain statistical properties and the volumes of data travelling across networks may be studied statistically. Over the last two decades much theoretical work has been carried out looking at these statistical processes [Akimaru et al, 1988], [Jain & Routhier, 1986], [Reiser, 1978], [Tobagi et al, 1978], [Schwartz, 1977].

For computer communication systems the traffic demands are usually expressed in terms of the call or packet arrival rate and the call/packet lengths. The statistical distributions of arrival rate and call/packet length have been monitored closely over many networks and found to conform closely to Poisson distribution arrivals and negative exponential distributed lengths.

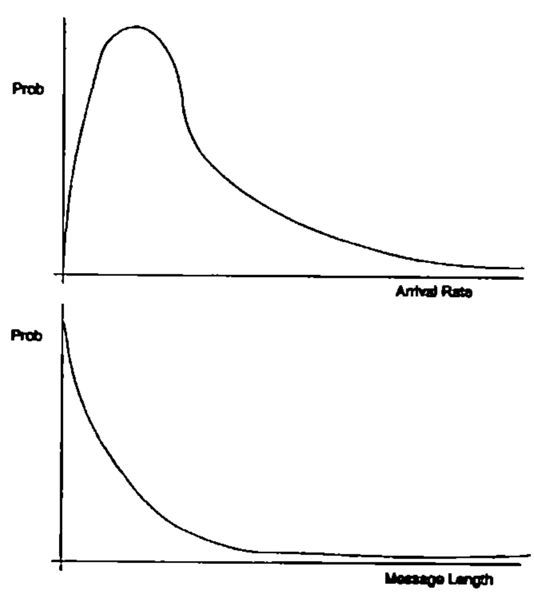


fig 2.13 The Poisson (top) & Negative Exponential (bottom) distributions

The Poisson distributed arrival rate process is the most typical statistical model of random time based arrivals because it represents a 'memoryless' system where the probability of an arrival in any one period is not dependent upon previous arrivals. The negative exponential message length distribution shows the way in which shortest messages are most likely with long messages all having a finite, but ever diminishing probability.

2.5.2. Virtual Circuit & Datagram Services

The circuit established in a data communications system may offer only a logical connection between the caller and the called, there are two distinct ways of providing such a service.

A virtual circuit provides a fixed path through the network for the duration of the call, unless a failure occurs. Either packet switched networks or TDM networks may use this system. The same path is used for transmission in each direction. Virtual circuits are defined as being either permanent (permanent virtual circuit = PVC), in which case they act in the manner of a fixed leased line or as being switched (switched virtual circuit = SVC) and they may be allocated and deallocated on request by the user. This is how the NET TDM system used by AT&T ISTEL to provide the core network operates.

The datagram system allows packets within a single session to take differing routes during a call and is described as a 'connectionless' service. The disadvantage of this system is that packets may arrive out of sequence and a method is required to re-order received packets where this occurs. Datagram services, being connectionless, cannot therefore have symmetric transmission paths. The AT&T ISTEL ADS packet network is a datagram service and operates using Bolt, Bereneck and Newman (BBN) packet switches.

In a datagram system, in order to send a stream of packets of data to a specific destination each one is addressed with the destination address, which represents a packet overhead. Each network node is then required to route all packets in the direction of their intended destination. Routing is usually performed along paths of minimum delay, the choice of path being dependent upon the network performance at the time of the packet forwarding. Routing selections are be applied independently for each packet and therefore dynamic use of network capacity in response to varying load may improve network delay performance.

Virtual circuit operation provides a continuous logical circuit where all packets arrive in sequence it is therefore necessary to inform each node along the selected path where to direct each packet. The call setup procedure may be complex and takes a finite time though the packet forwarding at each node does not require any route selection. Datagram systems require no call setup but do have the disadvantage that packets may arrive out of sequence.

The largest implementation of the datagram service is in the US Internet, followed by the Canadian Datapac network which uses datagram transmission internally but has virtual circuit protocols operating to interface to the user who 'sees' a complete VC system. Virtual circuit operation is more common and used by Tymnet, Telenet, BT PSS and the French Transpac.

2.5.3. New Packet Based Technology

Packet switching was developed in the 1970's when bit error rates on the lines were significant at the 10^{-5} level. This resulted in the implementation of many data overheads in the system to allow for recovery following transmission errors.

Advances in switching and processor technology coupled with the availability of near zero error rate trunks has meant that new packet based techniques are possible. New high quality data links in Europe and the USA have benefited from the advances in digital integrated circuits and error rates are now better than 10^{-11} and much of the error recovery of X.25 is largely a redundant overhead when used over such circuits.

Frame Relay [Schlar, 1990], [Gasparro, 1992] and Fast Packet [Degan et al 1989] are new high speed protocols designed for use on high reliability trunks (better than 10^{-10} error rate). Fast

Packet is a StrataCom Inc (USA) proprietary implementation of Frame Relay that offers a full Frame Relay interface but at the internodal trunk level operates using fixed length 53 byte 'cells'. This is essentially a slow speed (1.544/2Mbps) version of the Asynchronous Transfer Mode [McQuillan, 1991] cell system intended for trunk speeds at E3, Europe and T3, USA levels and beyond.

Frame relay is seen as a high speed replacement for X.25 in networks requiring the transmission of high throughput bursty traffic, such as that now found on Ethernet and Token Ring LANs serving IBM PCs. Minimal error checking is performed at intermediate nodes and any damaged frames are destroyed without notification, leaving the transport layer software at each end node to handle the retransmission on error and frame sequencing activities. The improved simplicity of the data link layer allows the internal network switches to operate at a faster rate and achieve much higher throughput than for X.25 operation. This is because none of the complex X.25 packet decomposition and analysis is required at intermediate nodes.

Frame relay operates with trunk speeds in the range 256kbps to 2Mbps. AT&T, the parent company of ISTEL, has a worldwide Frame Relay service based on Stratacom switches called InterSpan, for which the AT&T ISTEL TDM core network is used to carry the intra-European trunks. These therefore represent significant capacity demands upon the TDM core network.

2.6. Service Availability

The users of computer networks contractually agree to the cost of network services based upon service levels committed to by the network provider. The network provider must therefore determine the Service Availability, SA, to be offered. For the AT&T ISTEL TDM network the SA is a measure of network performance, (typically determined in hours of network availability per month) that the user can expect and the network provider must meet.

In telephone systems the equivalent is the grade of service, GOS, which refers to the probability that a call will be 'blocked', the network will have insufficient circuit capacity at a minimum of one access point through the system and will normally be 'lost'.

In the TDM network operated by AT&T ISTEL there is no equivalent of a 'blocking' GOS since all calls are based on a managed capacity allocation system. A network modelling system is employed to simulate the routing of all calls over primary paths and ensure that all calls survive the failure of any one trunk. The probability of trunk failure is the most significant reliability, and hence SA parameter.

PTT suppliers of leased lines in Europe set target levels of service, usually quoted as a percentage service availability [Data Communications International, Oct 1992]. Typical availability rates range from 99.4% to 99.7% over a 1 year period. It must be noted that these figures do not include planned outages (deliberate engineered trunk downtime) and are also targets, not guarantees.

The service availability requirements can have a major influence upon the network topology if trunks are provided by the PTTs with a lower availability than that required. For example, if PTTs provide 99.6% reliability for international trunks and a 99.6% availability is offered to the customer then the network topology will require primary paths of single trunks between customer sites, no duplicates being necessary since the circuit reliability meets the customer's requirement. Though, equally no multiple hop paths can be tolerated since the aggregate availability would drop. If PTT reliability is 98.5%, for instance, then there are two options. Firstly (and most practically) a lower grade of service might be offered to the customer. Alternatively, backup paths and additional capacity must be implemented to increase the network reliability to the level required by the customer. In order to meet a contractual service availability greater than that provided by the PTT there are, by necessity, increased costs related to the provision of additional, backup capacity. This increased cost is passed on to the customer and must therefore be commercially viable. The cost of providing increased reliability follows the so-called law of diminishing returns and it is therefore important to offer a grade of service compatible with economic reality. Figure 2.14 illustrates the way in which costs may gradually rise up to a point where single duplicate paths are made available, rapidly increasing as more and more alternate paths are necessary to offer reliability approaching 100%. An important limit on the economic sense of such increased expenditure is that the core network reliability should not be improved beyond that of the access network. There is no benefit to the customer in offering a highly resilient core network when the access network (or access 'tail circuit' to the customer premises) or access and customer hardware becomes the weakest link in the reliability chain. It is usual for network reliability to be expressed separately for the core and access networks since the access network reliability is largely in the hands of the PTT and should it be insufficient then measures can be taken. It is possible to improve access reliability by using dual access circuits to physically separate PTT exchanges.

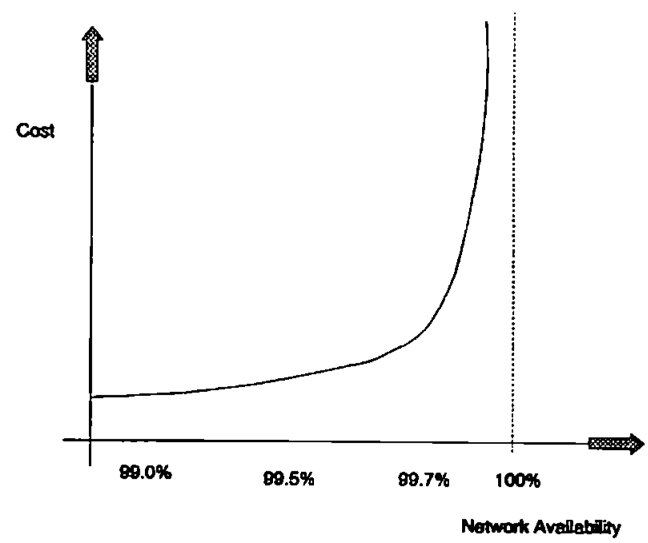


fig 2.14 The rapidly increasing costs of improving network availability

The grade of service is usually defined and bounded in product definition documents which are comprehensive and may run to a number of volumes [AT&T ISTEL, UK]. For packet based services the Grade of Service is quoted as a one way single character transit delay. This is a measure of the time taken for a single character from the user equipment to be passed through all intervening nodes and exit the network at the far end. The AT&T ISTEL packet switched service, ADS (Accunet Packet Service) is committed [AT&T ISTEL ADS Product Definition Document, volume 5] to providing a single character, one way, transit delay of 500mS through a maximum of 5 nodes within the UK or 750mS between differing countries in Europe. As a further example, the specification of the single character delay from the original ARPAnet design is 200mS [Frank & Chou 1974].

2.7. Routing

The route selection process in a network may be broadly classified in two ways. Firstly a distinction is made between systems that employ dynamic and static routing. Paths are either selected once for each end to end node pair, or they are recalculated at regular intervals to account for traffic fluctuations. Dynamic route selection is typically used in datagram type packet networks such as the Internet (ARPAnet) [Schwartz, 1977], in a TDM system it is not necessary to recalculate paths since the traffic is carried down fixed capacity channels.

The second distinction is between networks that use centralised route selection and those which use decentralised selection in which each node determines the routes to all destinations. The centralised route selection system uses a single processor to calculate all the required routes and sends the route data to all network nodes. This reduces the processing requirements at each network node. This system is not adopted in the AT&T networks since processing power is readily available in all network nodes at low cost. There are also considerable risks in relying on a single node to control the routing in all others.

The route selection is based on finding the lowest 'cost' option, where the cost may be measured financially, in terms of delay or in distance. The particular metric or combination of metrics for evaluating the 'cost' will depend upon the application. For example the AT&T ISTEL ADS packet network dynamically selects paths based on minimum delay since this is the performance measure most critical to customer applications.

The NET TDM network uses static routing based on minimum total trunk cost, where the cost of each trunk in the network is assigned by the network provider. It is possible to set the link costs proportional to the true financial cost of each trunk. In Europe, since the costs are dominated by the circuit setup charge and circuits are similarly priced for equal capacity around Europe, the link cost parameter is largely dependent upon the unit channel cost. This drops with increasing channel capacity and hence leads to routes being favoured where they use higher capacity trunks. This contrasts the link costs experienced in the USA which are largely distance related and leads to a need to make the 'link cost' trunk parameter reflect the true cost. If link costs are equal the selection of the minimum cost path then becomes one of selecting minimum 'hop' paths. This offers a means of controlling the route reliability since the service availability drops with increasing path length. It also minimises path delay since even though the NET IDNX is a TDM system there is a 4.1mS delay through each 'hop' which can be problematic to some customer services if more than four or five hops are encountered.

It is important to point out that for modelling purposes Gerla [1973] has shown that optimal static routing may be assumed in networks that use dynamic routing. Therefore, where a dynamic routing system is to be modelled it is possible to determine the optimal static routing and use this in its place, thereby simplifying the procedure. This is therefore relevant when modelling the AT&T ISTEL ADS network.

2.7.1. Routing Hop Limits

Each hop in a route is defined by the use of each trunk between end nodes. Routing systems are usually designed to minimise some form of cost function, typically involving physical cost and, often, a time delay term. For European international circuits the trunk setup costs dominate the overall cost and great importance is therefore attached to minimising the number of hops in a route.

There is a human expectation that minimum cost routes use the fewest hops but this is not the case in many network designs because cost savings are made where high capacity links are available and corresponding channel costs are lower. Where the distance related component of the link cost dominates it might be possible to use more longer links than the direct path to utilise the lower per unit distance costs. It may also be possible to use a number of short hops between nodes to form a route of lower overall cost than an equivalent length path of fewer hops.

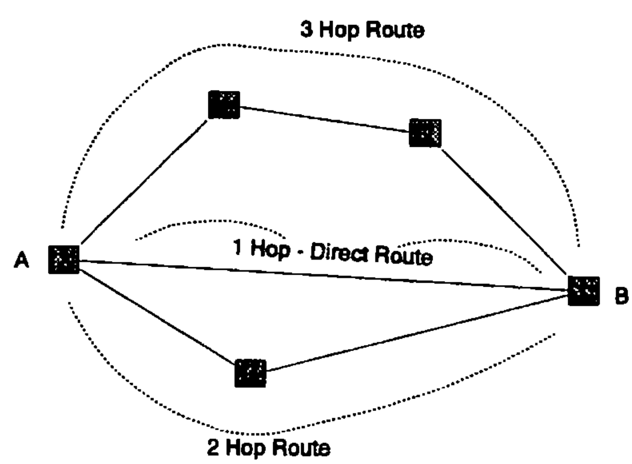


fig 2.15 Illustrating the hops between nodes compared to the direct route

In selecting suitable routes, a maximum number of hops is usually applied. This is primarily done to limit the number of possible routes to be found between a large umber of nodes but also forms a basis for ignoring high cost routes without needing to evaluate their cost directly. For the AT&T ISTEL ASDS-E network this limit is 12 hops.

The minimum number of hops in a route is desirable from a reliability perspective since the total route failure probability will increase with the number of route hops. Saksena [1989] includes the use of hop limitations in the route selection process of his network design work, but this is unusual, with Boyce, Fahri & Weischal [1973] and Gavish & Neuman [1989] relying on the minimum cost routing algorithm, ignoring the number of hops.

2.7.2. k-Connectivity

One of the common methods used for describing the routing capacity of a network is the '*k-connectivity*' term [Frank et al 1972], [Pierre & Hoang 1990] used because it is simple to evaluate in network modelling. It refers to the minimum number of links to which each node must be connected and is therefore evaluated by counting the minimum number of entries in each row of a connection matrix. A connection matrix indicates whether a path from each source to each destination node exists.

Source →

		1	2	3	4	5	6
D	1	0					
e	2	1	0				
s	3	1	0	0			
t	4	0	0	1	0		
↓	5	1	0	0	1	0	
	6	0	1	1	1	0	0

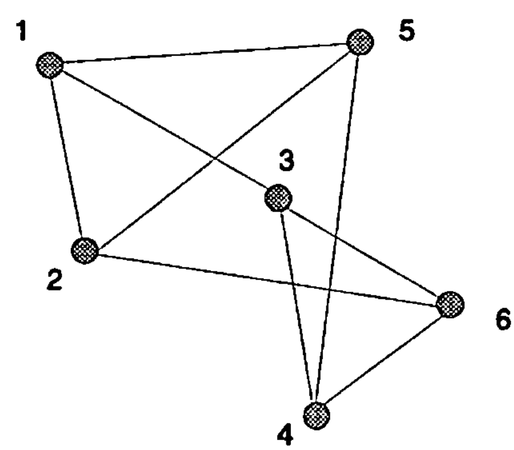


fig 2.16 A 3-connected topology

This is then a measure of the number of alternate paths a route might have by which to leave a node. Thus a 3-connected network guarantees 3 links impinging on each node, implying that a route entering on one link has a choice of two on which to leave. This method gives an approximate measure of the resilience of a network design and is used irrespective of the capacity allocated to the links. There is no guarantee that each outgoing link offers a distinct alternative path to any particular destination. It has been shown [Esfahanian & Hakim 1985] that it is possible to evaluate the maximum hop lengths of routes for a given k-connected network but the work is complex and a simple specification of the hop limit of a route would be more effective in a limiting a route search algorithm.

It is possible to specify a network with high node connectivity but with low capacities that cannot handle the volume of traffic required of them under certain failure conditions. The real problem with this system is that it can potentially over-specify the connectivity of a network and can waste link allocations. The main application of the k-connectivity measure is in the generation of initial solutions to network design problems where the quality of the result is not the main concern, merely the feasibility.

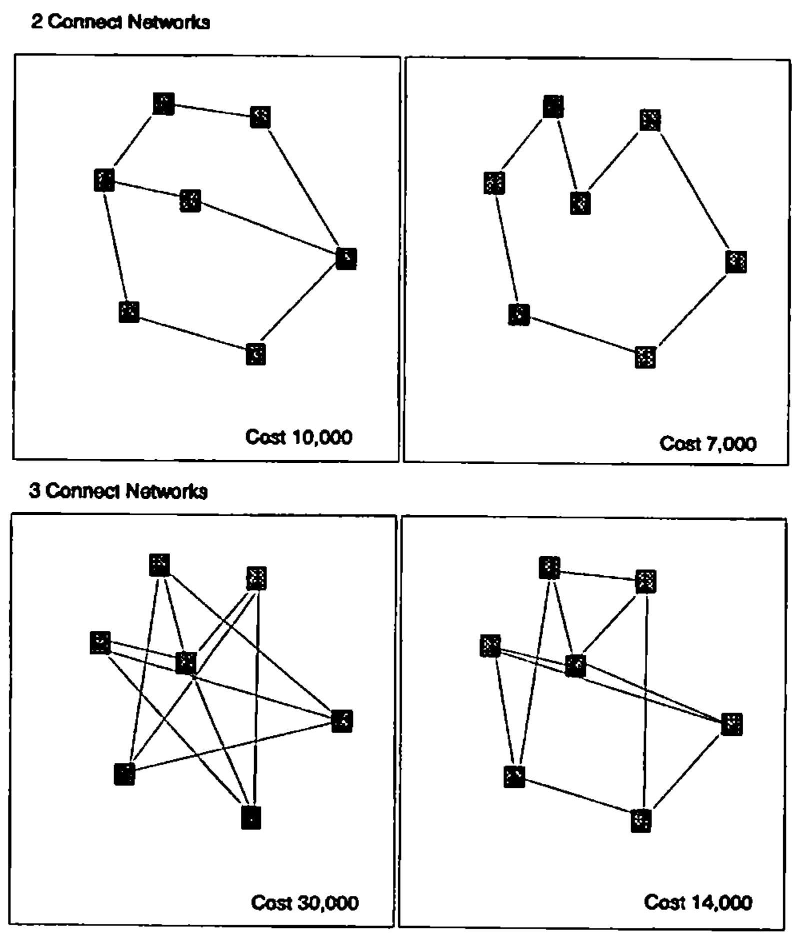


fig 2.17 Efficient (left) and inefficient (right) 2 and 3 connected graphs

The diagram above shows how it is possible to have very different networks, each satisfying the k-connectivity limit yet costing very different sums of money. While the k-connectivity constraint is imposed in a number of network design papers, it is used for the sake of simplicity in offering link failure resilience. As figure 2.17 shows, there can be great cost differences as a result of the connectivity limit yet in terms of resilience offered there are more issues involved. The connectivity limit is intended as a measure to ensure resilience, but the implication of this is that alternate paths must be available. It is quite conceivable that the k-connectivity constraint may be designed into a topology, yet in practice, for a given traffic matrix the resilience paths may be infeasible. This would require that more capacity is allocated. This is likely to be a consequence of the conflict between the use of k-connectivity and the analytic requirements of such a decision. The implication of k-connectivity is that the network will survive k-1 link failures. In order to allocate sufficient capacity to cater for this it is necessary to analyse the worst case effects of losing each and every permutation of k-1 simultaneous trunks. For a network of any typical size, say greater than 15 nodes, the loss of 2, 3 or even 4 trunks leads to a large number of permutations. This has therefore complicated an already difficult task. The simple k figure is used for a reason at odds with the true requirements of such a technique. K-connectivity is therefore viewed as being unsuitable for defining resilience needs and cannot be regarded as a viable component in quality design systems for reliable networks.

2.7.3. Allocation of Backup Routes Under Failure Conditions

In meeting agreed service availability levels the failures must be catered for by the provision of backup paths. The likely failures are that node device (card, processor) component will cease to function, typical causes being power failure, electrical failure, software error or physical damage. Link failures are commonly due to physical line damage (building worker damage and trawling of undersea cables), PTT equipment failure, accidental removal of cables/connectors or signal repeater power failure.

When a primary route fails for any reason it is essential that the network can offer a backup alternative route for the duration of the primary route failure. Because the two most likely causes of failure are node and link loss it is vital that the backup route uses no links or nodes in common with the primary route. This is usually described as a link/node disjoint backup primary route pair. Since it is not possible to have a node disjoint route pair that is not link disjoint this is usually just referred to as a node disjoint path pair.

In order to ensure that live network topologies are resilient to any single trunk failure, as specified in customer service contracts, AT&T ISTEL uses a modelling system (developed in-house and based on the AT&T IVIS, Witness simulation language). The model simulates the operation of the NET TDM network and measures the maximum capacity requirements upon the network following each possible trunk failure. No new capacity is allocated to customers until is has been shown by modelling that each possible trunk failure can be survived with no loss of customer traffic.

2.7.4. The Route Lockout Problem

One of the problems with route selection where alternate routes are required is that it is possible to have a primary route that 'locks out' a good backup route.

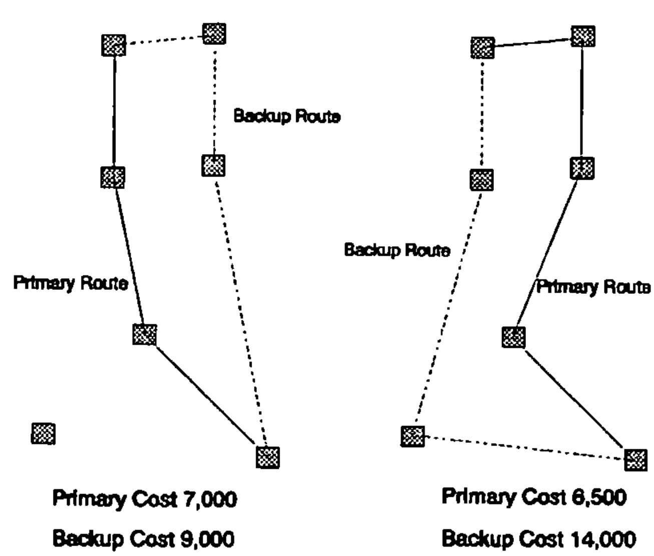


fig 2.18 An example of poor routes caused by primary route lockout

This occurs when a minimum cost primary route is selected which limits the choice of viable nodes to be used to construct an alternate disjoint backup route. The diagram above shows two possible primary paths, their costs being very similar. However, the resultant backup route choices have very different costs and it can be seen that the overall combined cost is considerable higher when the primary route 'locks out' a good backup.

2.8. Typical Access Network Configurations

The configuration of access networks is very much application specific and a large network offering a range of services may have many different equipment arrangements, but with common features. A range of access methods are available to the user such as modems, direct lines, and Packet Assembler Disassemblers (PADs) with the access method dictated by the applications used and any cost constraints.

In the context of large scale LAN based internetworks over a wide area a further distinct type of access network becomes apparent. The interconnection of remote LANs, using bridges or routers, across a wide area networks effectively makes the LAN the local access network and the bridge or router becomes the concentration point for the wide area.

2.8.1. Low Rate Asynchronous Terminal Access

The cheapest, most common access method used in large networks is the asynchronous (async) terminal. Characters may be sent and received at a fixed rate but with no timing constraints on when the transmissions may occur. Either dumb terminals, typified by the so-called glass teletype or the Viewdata terminal are most common. Viewdata terminals display text and data in colour with simple graphics characters and screen formatting commands in much the same way as Teletext on domestic televisions. Line rates typically range from 300bps to 1200bps and more recently moving up to 2400bps. The terminals are either directly connected to the local multiplexer using on site wiring or connected to a remote multiplexer using leased lines supplied by the local PTT. Alternatively, for locations that are very distant from a multiplexer or those with low budgets, the connection may be established using modems and 'dialling-up' access to the multiplexer. Terminal access is usually used to access remote host computers for tasks such as querying databases or viewing and depositing data on some form of booking or scheduling system.

The host processors may be buffered with front end processors to off-load some of the simpler tasks and reduce core network traffic. The access portion of the network might connect to a front end processor, often referred to as a 'menu node' because its primary function is to assist users with the logon sequence by offering a range of options in the form of a menu. Between the user and the menu node there may be more than one multiplexer, it is not uncommon to have two levels of multiplexing on a large access network. The AT&T ISTEL asynchronous (async) network is called Infotrac and offers nationwide access to mainframe host computers for mortgage, insurance and holiday reservation systems. In the Infotrac access network the user will typically connect to a small (8 input) local multiplexer on their premises, this in turn is connected to a larger multiplexer central to the district (typically 32 inputs) and then connected to a major node multiplexer for the region with up to 511 inputs. The major regional multiplexers are connected across the country by 'inter-nodal trunks', usually at a rate of 64kbps.

Low rate terminal access to mainframe systems can be provided through the use of front end processors (fep) where the fep is responsible for establishing communication sessions between the user and the mainframe application. This is particularly common in IBM mainframe systems running SNA (Systems Network Architecture), the IBM wide area networking system.

2.8.2. Packet Switched X.25 Access

Rather than use statistical multiplexers to offer multi-user access to the network, it is possible to use packet switching techniques to interface the users' async terminals to a PAD and onward into the network.

It is becoming increasingly common for users to require higher data rates than those usually offered by async terminals upwards of 1200bps, 2400bps, 9600bps or 19200bps. This is a result of new applications and the increasing volumes of data provided by the traditional async services such as holiday, mortgage and insurance quotations. There is a trend towards placing X.25 interface cards into PCs, allowing a number of users to access X.25 networks directly and benefit from the addressing features of X.25 which allow a range of different services to be selected by the user. The number of users that may be connected to an X.25 node will be determined by their particular traffic profiles and it is necessary to charge customers on a packet basis as opposed to a connect time basis which is common for async connections.

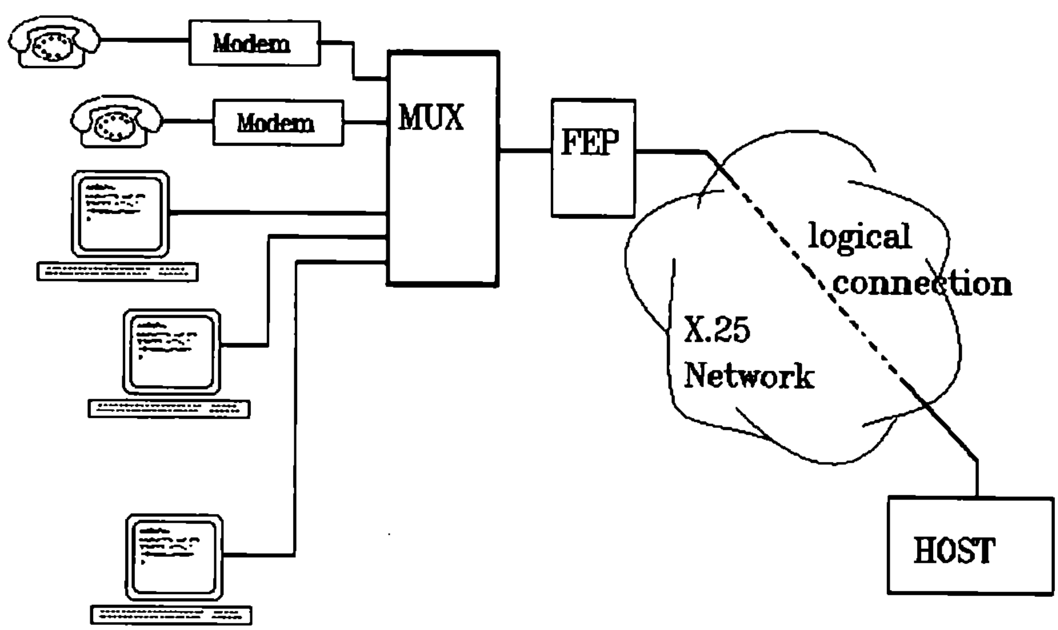


fig 2.19 A typical network using muxs, modems, feps, x.25 links and pads

2.8.3. Network to Network Access

Where a user has local access to one network and requires the use of another network, it is possible to offer direct network to network services. The advantages to the user ~~is~~ are ones of increased service choice and lower costs. If the network to which he connects can be accessed using local dial-up connections the phone charges are considerably less, by using the local network to access a distant one, than a long distance call to the second network. A further advantage of network to network connection is that it allows one network operator to offer a much wider range of services by encompassing those of third parties.

The connection of similar network technologies and protocols may be performed by direct connection whereas heterogeneous networks and protocols must be connected by gateways. The function of a gateway is to translate between the protocols operating on each network, it is responsible for establishing sessions and converting between the two datalink and network layer protocols. Some of the more typical applications of gateways are in offering connection to national and international packet networks such as PSS (UK), Datel (Germany), Internet (USA) and Transpac (France).

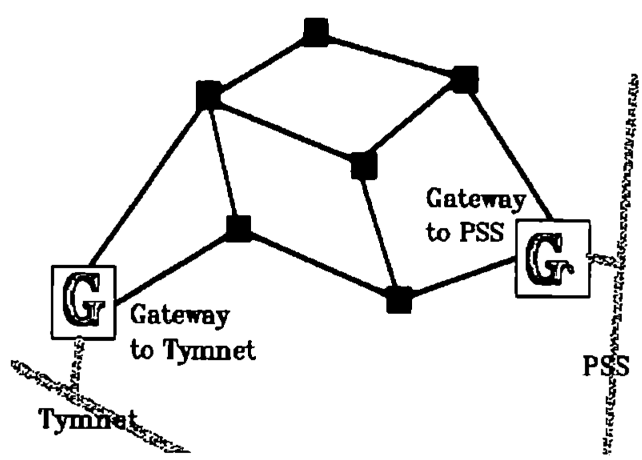


fig 2.20 A network with gateways

2.9. The ARPAnet

The first two years of growth in the APRA (Advanced Projects Research Agency, of the United States of America, Department of Defense) network topology is shown below to illustrate some of the basic features of a network topology and its evolution. The ARPAnet, as it was known, was originally started in the US in 1969 with just four nodes. The network has grown dramatically since its small beginnings and is now connected to every major international network.

The ARPA network has always been treated as a research project in its own right and for this reason there is a considerable volume of published academic work describing many aspects of its development [Frank, Frisch & Chou 1970], [Kleinrock, 1970], [Frank & Chou 1972], [Frank, Kahn & Kleinrock, 1972], [Frank & Chou 1974]. Diagrams of the ARPA net as it grew are shown by Frank & Chou [1974]. The configurations are interesting because they reveal a general aim of the design to incorporate a number of resilient loops. Some of the design features are relevant to later chapters of this thesis and some fundamental properties of the designs shown are:

there are no cross over links;

every node is connected by a minimum of 2 links, except in the initial topology;

no node is connected by more than 3 links.

When the network was small and reconfigured all existing links were retained. As it grew (from 24 nodes to 34) some links were discarded and replaced. This is significant because it shows that the design rationale changes as the scale of the economics grows, it also shows that the cost savings to be made by the trunk deletions were greater than the costs of the disconnection and setting up of the new trunks. As general principle it is cheaper to discard links and select new ones than it is to continue adding further, less than optimal links. Since the ARPA network is based in the US the cost penalties of ceasing trunks are small. The decision to cease trunks in Europe is more complex since international trunks are provided on one, two , three or five year contracts, the longer the contract the greater the discount; however, there is a significant cost penalty to be incurred if trunks are ceased. For this reason there is a considerable difference in green field and incremental design exercises.

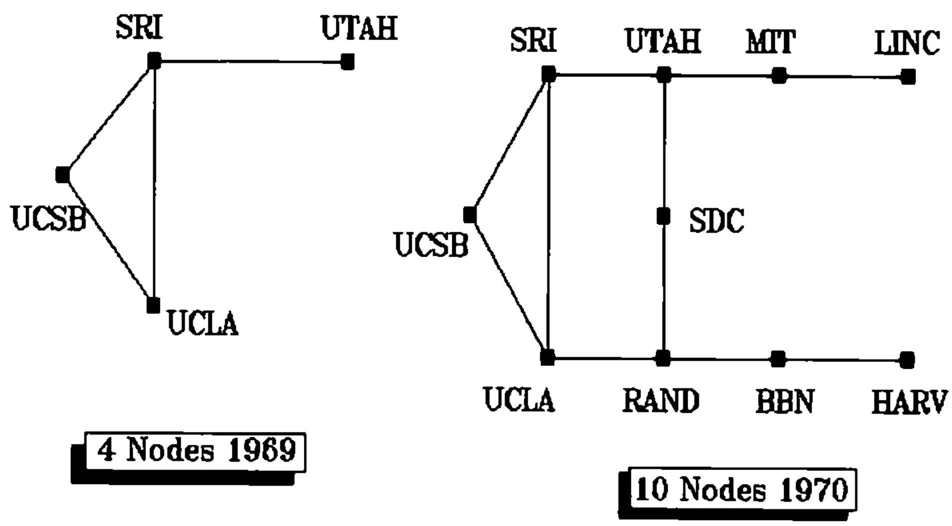


fig 2.22 The ARPA network in early 1969 and 1970

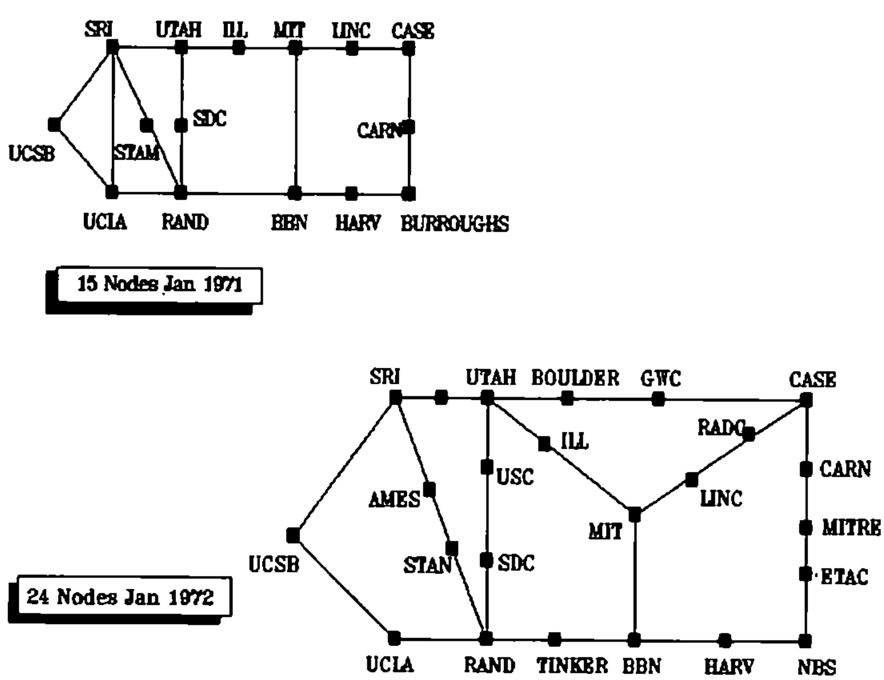


fig 2.23 The ARPA network in 1971 and early 1972

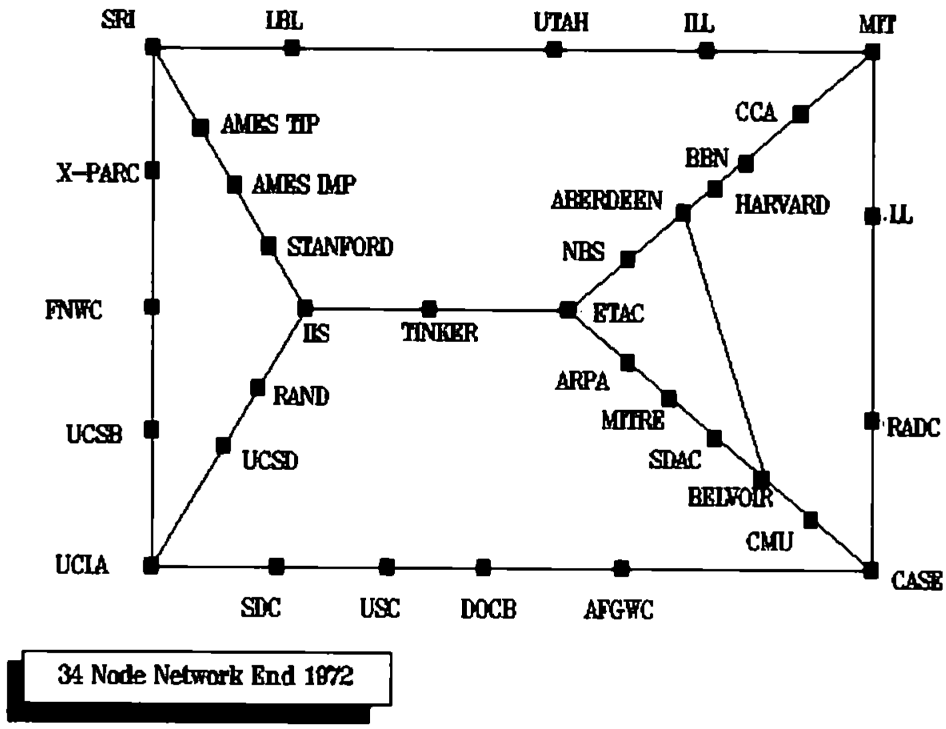


fig 2.24 The 34 node ARPA network from late 1972

2.10. Topology of the AT&T ISTEL Network

The AT&T ISTEL network design has evolved rapidly over the last few years since its inception in 1986 as part of the British Leyland (BL) car group's Data Communication Group. The evolution of the network has taken place in four distinct phases. The first phase was based around a UK-wide asynchronous terminal to mainframe host network using proprietary DCA (Digital Communications Associates Ltd, now part of the Racal group) equipment. AT&T ISTEL is based at the original BL computer data centre (CDC) in Redditch where a large number of IBM mainframes are securely sited. These mainframes now run a number of database systems for the insurance, holiday and financial industries, for which much of the AT&T ISTEL network still provides access.

The second phase involved the creation of a high capacity X.25 network around the UK to provide a resilient core network on which to support the DCA network. The requirement for a high level of service availability lead to the selection of BBN (Bolt, Beranek and Neuman) PSNs (packet switch nodes) which operate adaptive routing and can automatically deal with trunk failures to provide a reliable service.

The third stage of the network development was to replace the BBN core with a NET TDM network. The primary purpose of this was to offer an expanded range of network services, in addition to async and X.25 services the NET TDM allows the provision of dedicated 'clear channels' at port speeds from 300bps to 1536kbps. This allows the connection of a range of proprietary network systems across the network, that would not be possible with the DCA or BBN, X.25 network. In addition to the customer services the NET will support it is used as the underlying core network for carrying the DCA and BBN internodal trunks.

The fourth phase of the AT&T ISTEL network is currently operational and represents the expansion of the UK based NET network into mainland Europe. Much of the work detailed later in this thesis was used to design the European network which has annual trunk rental costs running into many millions of pounds.

2.10.1. Phase 1, Asynchronous DCA Based Network

The DCA network is based around access multiplexers in each major city and town in the UK. Users may gain entry to the network by two methods, either by dialling up modems at the access site or by direct connection using a PTT supplied leased line. Modem access is now available for approximately 98% of phone users in the UK at local rate. The DCA stat-muxs have 512 inputs (model DCA-375), 32 inputs (model DCA-120) and 8 inputs (model DCA-110). The 375 multiplexers act as the major nodes for the DCA network, between which are connected the internodal trunks. The 120 and 110 multiplexers are used purely as access devices.

The DCA-375s use DDCMP, the DEC communications protocol, which includes routing functions in addition to multiplexing. The DDCMP protocol is described in more detail in chapter four since it represents a small but significant increase in the traffic volume. One limitation of the DCA network is that its routing tables must be pre-configured by the network operator. This configuration is time-consuming and difficult to ensure that all routes are valid, and it becomes increasingly difficult to maintain as the network size of the network increases. There is also no guarantee that the routes selected by the operator are efficient and they are not automatically adjusted to cater for fluctuations in traffic.

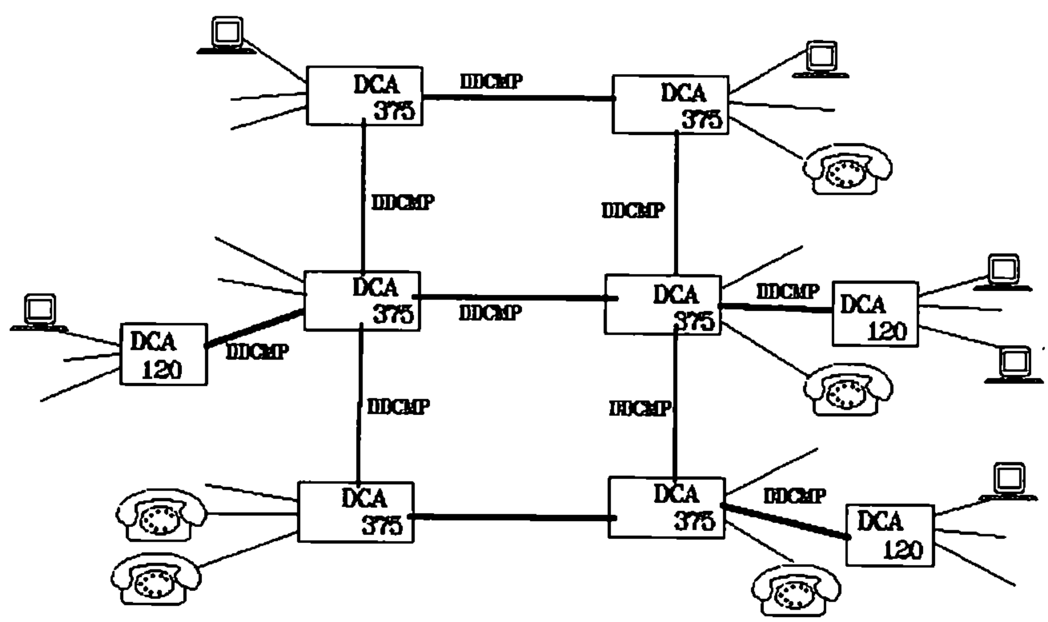


fig 2.25 The phase 1, AT&T ISTEL network, DCA based

2.10.2. Phase 2, the BBN X.25 Network

As the traffic demands on the original DCA-375 based network grew the need for a new high capacity core network was perceived. It was decided to implement an X.25 core network based on BBN C3 and C300 processors. This was intended to offer a resilient, configurable core network by which to provide internodal trunks for the DCA network in addition to offering a new X.25 service to customers. BBN are most well known for the design of the early ARPAnet integrated message processors, IMPs, [Frank & Chou 1974] which performed the major multiplexing and routing functions of the ARPAnet.

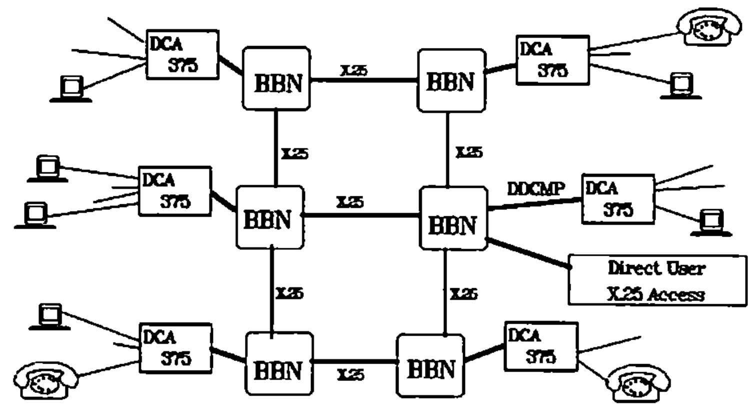


fig 2.26 The phase 2, AT&T ISTEL network, BBN based core with a DCA access network

The C3 processor has a typical capacity of 100 packets/sec and a cost (1993) of approximately £28,000. This processor can handle up to 512 input ports. The C300 processor has a capacity of typically 400 packets/sec and a cost (1993) of £80,000, and can handle up to 1200 input ports. The BBNs were sited around the country at the same locations as the DCA-375 multiplexers and used to form a lightly meshed network. The DCA equipment was then interfaced to the X.25 system to support the 375 to 375 links.

2.10.3. Phase 3, the NET TDM Network

As the traffic levels on the phase 2 network grew the X.25 network handled an increasingly heavy throughput. As the packet switch processor loading approached 80% moves were made to improve the core transmission capacity and reduce processor loading. Internodal capacity increases and a change in transmission technology were chosen.

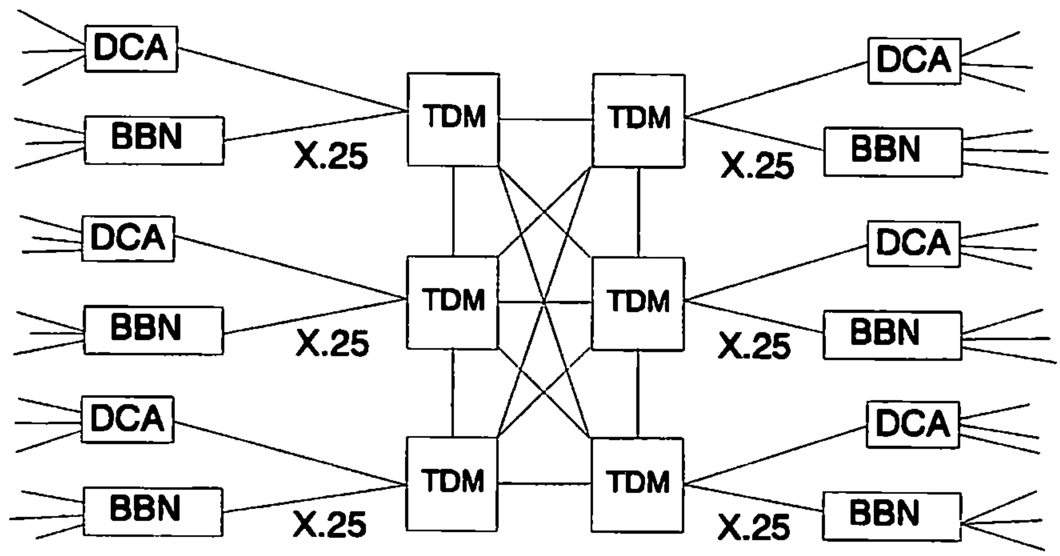


fig 2.27 The phase 3, AT&T ISTEL network, TDM core network supporting BBN and a DCA networks

The core network is now constructed from Time Division Multiplexers (TDMs) which can allocate capacity dynamically according to the offered load. The new TDM switches offer intelligent capacity management, being able to reroute calls around failed trunks and select optimal paths. Capacity is allocated contiguously with no fragmentation when low rate channels are carried (as would be a problem with a fixed time slot TDM)

Unlike packet switched systems there are no queuing time delay penalties to be suffered the only delays are due to trunk buffers which account for 4mS per trunk. Maximum efficiency is achieved when all inputs consume all available transmission capacity. Time delays are fixed, which greatly assists modelling of TDM systems. The major problem of TDM (and one that is possible to avoid with correct capacity allocation) is that if more transmission capacity is required than is available a call will be rejected. There is no degradation in delays with increasing throughput as is the case with packet switching, but the rejection of a blocked call remains a potential problem where capacity is limited. This is most likely to occur following a trunk failure when lost calls will be rerouted if possible.

TDM systems can offer a cost effective method for providing core networks, typical node processor costs might be £60,000 (1993) for approximately ten times the throughput of a BBN C300 switch. Each TDM has inputs made up of a mixture of X.25 outputs from the BBN equipment, DCA-375 HT protocol outputs and direct links from customer terminal equipment. The TDM is also used to provide customer circuits in the form of dedicated capacity, known as 'clear channel', these channels can carry any data required by the customer acting as pure unformatted data channels. The TDM network offers a complete end to end managed network for all customer circuits. It has advanced fault detection and reporting facilities which are essential for providing a fully managed service to the customer.

2.10.4. Phase 4, the European NET TDM Network

With the successful deployment of the NET TDM network in the UK it was then expanded to cover the major countries of continental Europe with a resilient mesh. The majority of internodal trunks are supplied at 2Mbps and the remainder are at 512kbps to countries with lower traffic requirements. The core of the current AT&T ISTEL network consists of approximately (1993) 53 NET, TDM facilities located throughout 15 European countries. AT&T bought ISTEL just before the expansion into Europe and the European NET TDM network is known as ASDS-E (Accunet Spectrum of Digital Services) which is the 'product definition' used by AT&T to describe the type of network. This reflects the fact that ASDS-E can carry a wide variety of different data services. Since the European expansion the original DCA and BBN services have been augmented by the provision of InterSpan Frame Relay internodal trunks and international, inter-LAN connections.

An important advantage of the NET TDM is that circuits may be reconfigured between different end-points from a single management node, this allows for changes in customer requirements to be far more rapidly met than might be possible using only dedicated circuits provided by a number of PTTs around Europe.

CHAPTER 3

Access Network Analysis

3.0. Introduction

The general remit of this work has been to analyse the issues that affect network performance and cost. It was indicated in chapter 1 that a differentiation is made between the access portion of a network that gathers all inputs traffic streams and the core network structure that transports the traffic across significant distances. This distinction is made in order that the functionally independent factors may be separated. The result of this is that the access network may be analyzed and a large number of inputs to each major node in the core network may be summed to produce a single load description. This greatly simplifies the analysis of the core network by reducing a very great number of data inputs, perhaps many thousand for a national network provider, to a single input traffic load value for each node. It might be argued that without such rationalisation of the problem there would be little chance of finding a solution, yet, with careful analysis there is little to be lost by performing such a simplification.

The basis of this chapter is the analysis of the access portion of a large network, such as the AT&T ISTEL asynchronous terminal based Infotrac service. The aim being to develop a method for finding the channel capacity required to provide a given level of service to the customer, measured in response time for interactive sessions. Traffic profiles from a large number of network users are used to create composite arrival rate and message length distributions. An estimate of the performance of a multiplexer arrangement with these composite traffic inputs is then developed. A single figure statement of the mean queuing delay does not provide sufficient indication of true network performance. A low mean delay might not reveal that 80% of delays are, for instance, ten times worse. Service contracts will usually be stated in terms of the, say, 90th percentile delay. It is therefore necessary to develop a means for estimating the probability of the delay performance meeting certain thresholds and presenting the figures in a simple fashion for every-day use by a commercial network provider.

3.1. Access Network Traffic Analysis

The access network is defined as being the multiplexer and data concentration equipment that presents the total input traffic to the core network for cross country transmission.

The AT&T ISTEL Infotrac operates what is termed 'host echo' where the host computer returns all characters typed by the user which are then displayed on their terminal. The alternative is 'pad echo' whereby each character is returned from the pad (or multiplexer) to which the terminal is locally connected. This has the advantage of avoiding all the internal network delays but the disadvantage that the character is assumed to reach the host correctly, error detection and handling procedures therefore become more complex.

The performance of the access network is seen by the user in terms of response times for character echo and traffic throughput. The throughput is the rate at which data can be transferred and is therefore limited by the line capacity and the method of transmission. The communication through the multiplexers uses a data link protocol which adds an overhead to the actual data transmission, this therefore limits the maximum possible throughput to less than the line capacity in the circuit. The time delay seen by the user, between pressing a key and seeing it echoed on the screen, is a function of the total number of users connected to the multiplexers, the statistics of the traffic they each submit and the capacities on intervening links.

The inputs to the access network design problem may be defined in two parts, one part is in determining the quality of service, typically response times and throughput, the user may expect and secondly deciding the location of multiplexer equipment in order to minimise line rental costs between users, multiplexers and the core network.

It is important to quantify the grade of service a network user receives. This is to ensure that the customer is satisfied with the network performance and also to ensure that the network configuration is efficient (minimum cost to the provider) for the load presented. A great deal of work has been conducted [Bahl & Tang, 1972], [McGregor & Shen, 1977], [Kleinrock & Kamoun, 1980], [Grout, Sanders & Stockel, 1988] in the field of optimal concentrator location for access networks. From this point on, the term multiplexer will be used to indicate a device that acts as data concentrator, protocol generator and communications processor. The location analysis is of partial use to a medium sized network operator since the optimal multiplexer positions are necessarily modified by external factors such as physical constraints, for instance finding room to house a multiplexer and providing 24 hour maintenance access. The calculation of the 'best' locations remains useful because of its possible influence on location choices, however it does not provide absolute answers. Due to these external factors and historical development in the AT&T ISTEL network most multiplexer locations are largely predetermined. The design problem then becomes one of determining which users should be offered access to each multiplexer, and what is the probable grade of service they might expect, for a given configuration. The access to particular multiplexers can be controlled by informing different groups of users of different telephone numbers where the multiplexers are operated in a 'dial-up' mode. Alternatively the connections may be made by connecting each user's equipment to multiplexers with leased lines. The final choice is being influenced by both cost and the expected loading conditions on the each multiplexer.

The AT&T ISTEL network customers may be broadly classed according to 6 major business sectors: Leisure (travel industry); Motor (car manufacturer); Finance (insurance and banking); Retail (shopping outlets); ISTEL (test and monitoring data) and Other (the rest). The classification by market sector also serves to typify the users by the statistical traffic profile of a network access session. As an example, the Finance users tend to have long sessions of query based calls as they interrogate a series of databases for insurance quotes. On the other hand the Motor industry users have sessions that are either purely stock entry or short duration stock enquiry calls. The traffic profiles of the data for user to host and host to user are therefore very different for the different users on a class by class basis but are clearly likely to be correlated within the classes due to the consistency of the data formats and query constructs.

The aim of analysing the access network traffic is to optimise the mix of users at each access point and to provide some estimate of expected resultant mean time delay. In order to perform such analysis a good starting point is to look at the reasons for data delays and determine what the network user 'sees' as a response delay.

3.2. Traffic Delay as Perceived by the User

Though the network user will see delays in the on-screen data appearing, due to the speed of the mainframe application, it is the response time for the mainframe/network to echo each typed character that is the prime performance metric. The mainframe response to a single typed character is assumed to be small when compared with the network delay, delays due to the application, such as a database are outside the control of the network provider and addressed separately.

Before looking at the methods for analysing traffic it is important to determine at what point in the network the major contributions to echo delay occur. The step by step analysis of the mechanics of character echo shows there is one dominant point of delay.

Consider the chain of events following the user typing a single key to select an option from a menu already on the screen. The sequence of events is as follows:

1. the character leaves the terminal and enters the multiplexer;
2. the processor in the multiplexer then buffers it along with all the other keyed in characters from other terminals connected to that multiplexer. Each character is also appended with an address (a protocol specific address marker) to indicate to the system from which terminal it originates;
3. the assembled characters plus the source address are placed in a data frame. The frame

will have a maximum allowable size and when enough data is present to fill the frame it is sent. If the frame is not full but a predetermined period has elapsed since the last frame was transmitted it is sent. The frame has a separate address added to indicate to the network from which multiplexer it originated;

4. the entire frame is transmitted through the core network to each node/multiplexer in a path until is reaches the node connected to the required destination host computer;
5. at each point where the original character is passed through an intermediate node queuing may take place;
6. the host computer receives the character and returns it to the terminal via its local multiplexer and all those intervening in the core network path, using the same frame formation process to indicate its arrival;
7. on its way back to the originating terminal the echoed character may suffer queuing at each intermediate node in the core network;
8. the echoed character must now return to the originating terminal, through the local multiplexer, for display on the screen. The echoed character must be queued in the local multiplexer along with all the other characters for each and every other connected terminal.

3.2.1. Terminal Traffic

In examining the system of character echo, above, one may note that the volume of data on the path into the network (terminal to host) is low because it is due entirely to keypresses from users to select menu options or type passwords/names etc. This data is therefore entered at human speed (perhaps between 1 and 8 characters per second) and represents a very small traffic demand on capacity.

However, the return path (from host to user terminal) is the main traffic carrier. It transfers all the 'screens of characters' that represent the information from the host's menus, prompts and application data. A typical screen of data with a banner heading and a range of application details might consist of between 100 and 400 characters and these are transmitted from the host at the maximum possible rate.

These screens of characters represent heavy, constant use of the data link and a number of such simultaneous transmissions could result in more demand on the system than the available capacity can satisfy. In a system designed to be highly cost effective, such as in a large scale asynchronous network like Infotrac, it is necessary to maximise the mean trunk utilisation for a given mean delay response. For short periods of heavy demand, following simultaneous transmission of full screens, for a large number of users, it is likely that higher response times may be tolerated while remaining within mean response time targets.

What is seen from this analysis is that the delay each single keypress experiences is very small on the path through the network to the host because the volume of traffic in that direction is small. This is true where all the calls are of an interactive type, delays would be very much more significant in the user-host path if larger traffic bursts were present, for instance due to the use of file transfer processes. Within the core network the queuing delays are likely to be small due to the high speed internodal trunks, typically 64kbps, which are monitored to ensure core delays remain low. It is the local multiplexer with many low data rate inputs and very often a low capacity link (typically 9.6kbps) to the core node that can result in some queuing problems if too many users with heavy demands are configured to use the same multiplexer.

A survey of network users was conducted by the author which revealed that the users' natural perception of the character echo delay is entirely due to what is considered to be the character taking a long time in being serviced by the host - it is believed that this is because the user has the misconception of a busy computer at the other end and a fast network on the host to user path. The user appears to perceive the computer being slow to respond to the keypresses and that the data, when serviced by the host, is transmitted almost instantaneously back through the network.

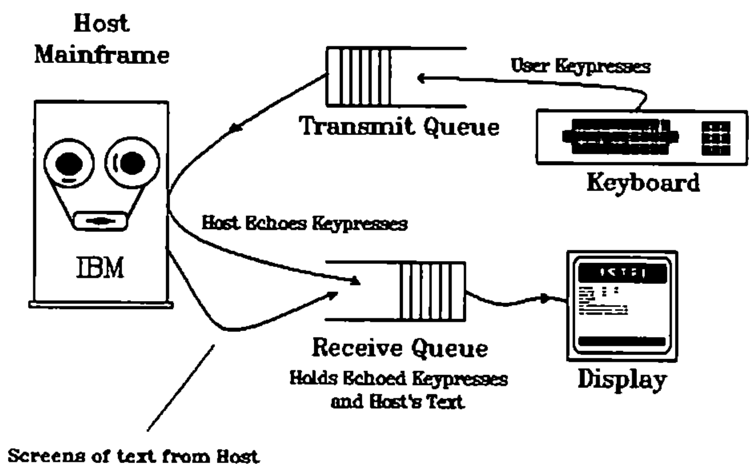


fig 3.1 The remote echo principle

3.3. Discrete Time Analysis

Discrete time analysis uses mathematical and computer models to represent the state of a system at any one time. Using a series of rules that mimic the real world the system then modifies the model's state in a periodic way. Very complex models have been developed to analyse the operation of communication systems. These systems were originally developed in the early 1970's, [Kleinrock 1970], requiring large minicomputers to simulate the many intricacies of the packet switch protocols. The latest systems now run on personal workstations, such as BONeS (Comdisco Inc, sales literature, 1993) which is a Block Orientated Network Simulation systems and the CACI ComNet Network simulator (CACI Ltd, sales literature, 1993). AT&T ISTEL uses the AT&T ISTEL IVIS (Istel Visual Interactive Services) Witness simulation package to model expected response times of X.25 switch arrangements in creating performance figure targets for customer contracts. It is this type of analysis that is best suited to very complex systems because it can mimic many complex state dependent events.

This type of analysis requires models that accurately reflect the state of the system under scrutiny. This usually entails computer data structures that hold the vital parameters describing each 'state' of each element of the model. A series of rules are then invoked to systematically transform the models according to the sequencing rules of the protocols.

A problem with such simulation packages is that the results produced are only valid for the individual inputs given. It is therefore necessary to perform a number of tests with slightly differing inputs to analyse the sensitivity of the model to its inputs and hence evaluate the quality of the results. This can produce long analysis procedures and it may be difficult to interpret the results in order to apply them to a practical application. Where simulation methods are used within AT&T ISTEL they are used to model only specific subsections of the network. Care is required to isolate only those areas of the network pertinent to a particular customer or project, since if the model becomes too large steady state results are not achievable within acceptable timescales, i.e. periods of less than a day or two.

Demonstrations of the Bones system were viewed but they proved not to offer the ability to reflect grade of service figures for various multiplexer loading configurations, taking a number of hours to produce figures for a small number of network queues. They were dismissed as unsuitable since the figures produced represented one off model events and it is not possible to evaluate quickly the sensitivity of the delays to changes in traffic and network parameters. In general, simulation systems require a compromise between simulation time and model complexity.

3.4. Probabilistic Analysis

A probabilistic analysis aims to use a series of statistical representations of various inputs, combine them in a manner that emulates the system being modelled and provide an output that describes the probability of each potential outcome. The advantage is that this system gives an overall 'most likely' result. Thus the entire operation of a multiplexer system can be modelled and used to determine its 'most likely' performance for a given set of inputs.

Alternatively one may simplify the results and express a single confidence level in performance reaching or exceeding certain thresholds. A result in the form 'the performance is expected to better a 500mSec delay for 95% of the time' is sought. This type of figure then gives the network designer a fast way of evaluating the relative performance of different multiplexer configurations.

Certain software packages of this type are in existence [Sauer 1978], [Whitt 1983] but they are of sufficient commercial sensitivity that full details are not released and there is no academic access to descriptions of methods employed. However Yokoyama, Miyake & Nakajima [1988] give an outline of their load dependent network analysis and Jain & Routhier [1986] discuss modification to the modelling to account for the so-called packet-train', where packets follow each other according to the probability of a session being in progress.

It is necessary to evaluate the various ways of expressing computer network traffic and determine the most efficient ways in which to present and analyse the data. Since there are many types of network user and many possible access configurations a method for modelling the general effects of various traffic levels on differing network configurations is sought. NETAL is a network analysis program, written for this project by the author, to analyse the total traffic loads and expected queuing delay for various multiplexer arrangements and customer traffic profiles.

The combining of the traffic profiles of a number of users from different market segments takes account of:

i. the type of user connected to each input of the multiplexer;

ii. the data transmission protocol used by the network.

The network configuration is based on the access class of each user connected to the multiplexer (e.g. leisure, motor, finance, etc.) because this determines the characteristic means of message arrival rate and length. The data transmission protocol is also modelled and the extra line capacity it uses is accounted for in delay calculations.

The results of this delay modelling are required in a form that is readily understandable by the network designer. Having calculated the range of possible delay values for all traffic levels NETAL presents the figures in graphs showing the expected delays for all possible arrival rates and message lengths, and then the probability of the delays exceeding a range of thresholds. i.e. the figures are expressed as a delay of τ mS being expected for better than θ% of the time.

3.5. Analysis of AT&T ISTEL Network Traffic

In order that the designers of the access network for the AT&T ISTEL UK network be able to estimate the likely delay figures for various user to multiplexer configurations, a series experiments were conducted with different models. These models take estimates of traffic statistics for each of the user classes and combine them to show the expected traffic delay probabilities.

3.5.1. Cyclic Traffic demand

The traffic level across a network is determined by the number of user connections required and is found to be heavily influenced by cyclical demand patterns such as seasonal trade variations, time of day variations and economic climate variations. Sharma [1988] recommends that periodic patterns to demand should be closely monitored due to their use in predicting future traffic demands and growth trends.

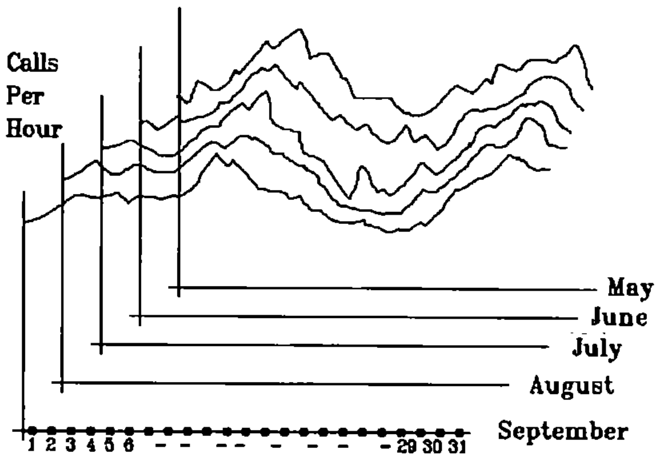


fig 3.2 Cyclic traffic demand over a monthly period

Not only may overall traffic demand profiles be cyclic over weekly, monthly or annual periods, but there may be market sector specific cyclic demands. For instance financial based services, such as investment and tax assessment services may be highly seasonal in relation to the financial year. Leisure based traffic demands may be cyclic in relation to school holidays and periods of fair weather. For large network providers there is great potential for understanding customer demand based on traffic monitoring by market sector. Where load grows year on year, the cyclic peaks may be monitored to ensure that provisioning of new capacity is made ahead of expected peaks due to both organic business growth and cyclic demand profiles.

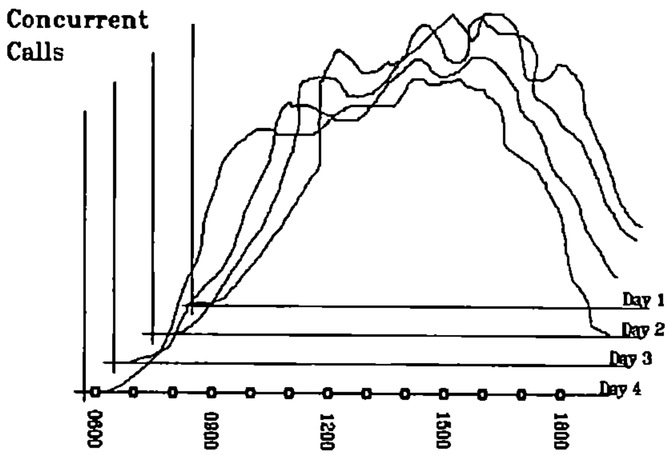


fig 3.3 Cyclic traffic levels over a daily period

Though the cyclic nature of the traffic demand is an important factor in determining network performance targets it would be prohibitively expensive to build a network to provide a constant or high level of service throughout the periods of heaviest traffic. For this reason it is normal to use the 90th percentile of maximum traffic level as a planning guide. For the purposes of performance modelling it is necessary to work on traffic figures for the busiest hour of the day, therefore the 90th percentile of the busiest day is selected. Where seasonal traffic variations lead to short term traffic peaks the network is operated under relaxed constraints, a poorer delay figure is accepted.

3.5.2. Call Arrival Rate

The call arrival rate distribution used in this work is that of the Poisson distribution. This distribution has been widely used in the modelling of arrival rates [Whitt 1983], [Ozekici 1990] and is generally a good approximation of network traffic call arrivals. Some preliminary work using traffic monitoring equipment on the AT&T ISTEL network has shown arrivals to correspond to this distribution. The mean and variance of the Poisson distribution are equal, represented by a single parameter, λ, indicating the mean arrival rate of traffic bursts per second. The probability of individual values of x being given by;

$$P(x) = \frac{\lambda^x . e^{-\lambda}}{x!}$$

The probability distribution function, pdf, of the arrival rate is shown below.

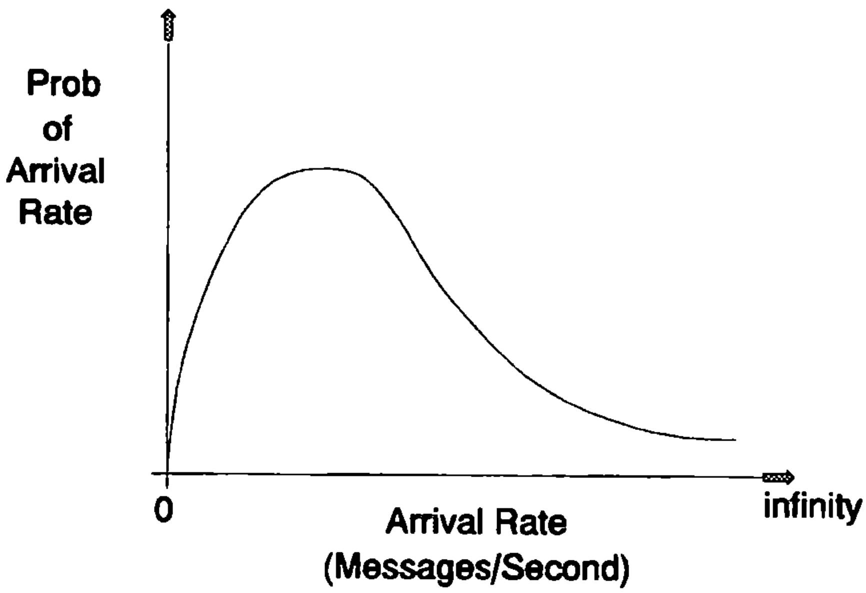


fig 3.4 The Poisson arrival rate distribution

One implication of this distribution is that there is no knowledge gained about the probability of a future arrival given that one arrival has just occurred. The Poisson distribution is therefore described as representing a 'memoryless' system. This is in contrast to the work of Jain and Routhier [1986] who attempt to show that in modern packet switched networks the probability of receiving a packet of data is dependent upon having just received a prior one. This uses the concept of 'packet trains' where the existence of user sessions is recognised as leading to a stream of packets from a user conducting a call 'session' rather than all packets having simple Poisson arrival rates. This is seen as being a second order arrival process where one arrival rate determines the operation of the second. The disadvantage of this is that it adds more complexity to the analysis for little discernable gain.

3.5.3. Call Message Length

Analysis by Grout [AT&T ISTEL, Call Profile Report 1989], [AT&T ISTEL Traffic Analysis Group, 1989] has shown that the distribution of call lengths closely follows a negative exponential distribution and the call messages are approximately negative exponential in distribution. Important features of the negative exponential distribution are the greatest probability of short messages and the probability of messages of increasingly long length not being zero, but diminishingly small as the length tends to infinity.

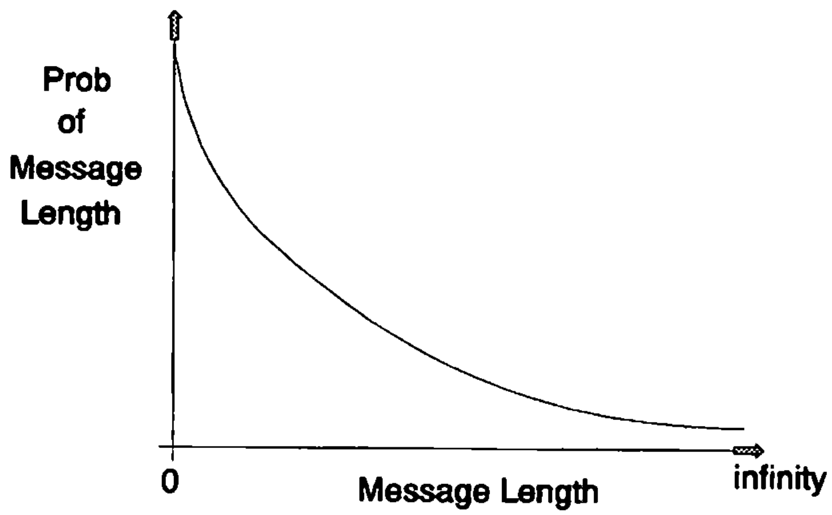


fig 3.5 The negative exponential message length distribution

3.6. Statistical Modelling of Composite Traffic

It is not a simple task to monitor the traffic statistics for a number of users because the process requires either a dedicated line monitor to record all traffic details or network management systems must be configured to log all traffic. Each process is costly in terms of time, equipment, transmission resource and analysis effort. However, some figures were taken from a number of traffic logging sessions and these were used as inputs to this work.

With the traffic on the AT&T ISTEL network shown to conform to Poisson arrival rate and negative exponential message length, a standard expression for estimating queuing delay may be used, this is detailed shortly in section 8.3. The model of a multiplexer and the users connected is used to create a combined traffic profile and generate performance figures. The NETAL package that perform this must take into account:

i. the desired format of results;

ii. the single user traffic model for each service class;

iii. the method for combining single models;

iv. the data protocol used on the network (this leads to extra data);

v. the computer processing power and memory available.

Project funding provided an IBM PS2 personal computer that has a 16MHz 80386 microprocessor and 640Kbyte of memory. The entire memory was required to hold all program code as well as program data which posed some limitations as to the size of model and volume of statistical data that was processed.

3.6.1. A Single Network User Model

In order to form a basis for a computer statistical model it was necessary to generate and store traffic profiles for each user class. Each user traffic model is made up of a fifty point sample of the call arrival rate and message length probability distributions. Fifty sample points are used due to the processing and memory limitations, to use more points greatly increase the calculation times for a complete analysis. The number of calculations is roughly proportional to the square of the number of samples. Fifty points were chosen to produce final calculations within a few seconds, rather than a few minutes.

A separate traffic profile, message length and arrival rate pdf is stored for each of the six different network user classes. The correlation of traffic profiles for each user within each class on the AT&T ISTEL Network is particularly strong and therefore seen as a solid basis for using a single model per class.

3.6.1.1. Leisure User Traffic

Within the travel industry the data transfer session is typified by the arrival of a customer in a travel agency and the initiation of an interactive session using a viewdata terminal contacting a remote host computer. The customer, with the travel agent, will view some holiday options, request further details of a chosen holiday or a small number of holidays and then either terminate the session or make a booking.

The traffic produced by such Leisure users during a session can therefore be broadly categorised by the following characteristics:

i. a burst of data to access the system and call up an index page for holiday information;

ii. a pause of a minute or two to peruse the display followed by a small series of short bursts of data to select specific information;

iii. a repeat of a large burst of data from the remote host to display another holiday option and a repeat of this burst/pause cycle until the user has found the desired details;

iv. a possible booking phase to confirm customer details - consisting of a stream of questions from the host and a series of responses by travel agent to give customer name and payment details;

v. session termination.

It is thought that the consistency of the traffic profiles is due to the highly trained operators who are familiar with the systems and have relatively uniform response times and can guide their customers to decisions within approximately constant time-frames.

3.6.1.2. Motor Industry User Traffic

A traffic user profile for motor industry customers is very different from that of leisure industry users, being typified by either short single part availability/pricing query sessions or by long sessions to enter large amounts of stock ordering information. The motor user session traffic is therefore likely to have the following characteristics:

i. a short logon period to select the desired host system and supply local user details;

ii. long repetitive bursts of data from user to host to request car part orders or availability and pricing;

iii. bursts of host to user responses to confirm entries.

The fact that the traffic bursts and the mean periods between them are highly correlated for the motor users is perhaps an indication of the consistency of the data entry procedures and a result of the users being familiar with the system. It is therefore not surprising that the traffic profiles can be modelled successfully.

3.6.1.3. AT&T ISTEL Traffic - from Real Traffic Capture

Some results from AT&T ISTEL's own traffic logging system are now presented. Due to the complexity and data overhead of the procedures required to gain this information it is only available for a single example. For every character that was generated the network created a logging packet that recorded the system time. The logging packets were then processed to give a plot of the character ratio throughout the session. The traffic on a typical leisure user's terminal was monitored for a small number of call sessions (period from logon to logoff).

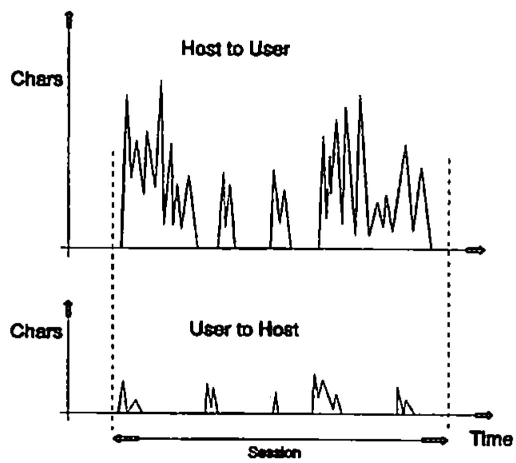


fig 3.6 Typical session traffic

From figure 3.6, above, it may be seen that the user-host traffic is all but insignificant compared to that for the host-user. However the traffic delay experienced by both types of traffic is very similar because the dominant cause of the delay is found to be the queuing on the return path. The profile of a typical session shows the user to type a few characters, receive a burst of data from the host, pause while examining the data and then repeat the process until the session is closed.

The distribution of the message lengths is approximately negative exponential. As expected for a statistically small sample there is some deviation from the ideal distribution, but if the sample size were increased then the profile is expected to match the theoretical curve more closely.

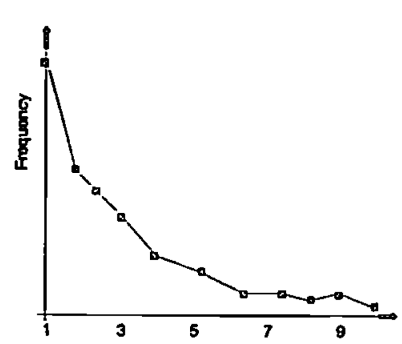


fig 3.7 Message length frequency gathered from logging data

The statistical summary of the logging exercise is shown in table 3.1.

Total Period Analyzed	3612 sec	
Session Time	3394 sec	100%
Non Host-User traffic time	2816 sec	83%
Host-User traffic time	578 sec	17%
Host-User bytes in total	54646	
Host-User bytes/sec	945	(equal to 78.8% of 1200 bps)
No of sessions	11	
Host-User bytes/session	4968	
Host-User traffic secs/session	53	
Average session time	309 sec	

Table 3.1 Summary of data logging exercise

These figures are interesting in that they show the host to user traffic is transmitted at the rate of 78.8% of the maximum line speed 1200 bps (945/1200 = 78.8%). The 21.2% of the data rate remaining was that required by the system for the Digital Data Communications Message Protocol (DDCMP, proprietary to DEC) used between the multiplexers. This is a significant factor and should not be ignored in delay calculations. The average session time of five minutes is also of interest because it points to the 'characteristic leisure call'. AT&T Istel managers have indicated that this figure is precisely what they expected from their own practical experience. They also confirm that using a small number of user profiles to characterise a range of users for modelling purposes is a reasonable approach given the consistency of sessions across the industries. This confidence in modelling traffic profiles is largely based on the training of the operator resulting in constant session activity and the similarity of the session types within each commercial sector.

3.7. Large Scale Model Construction

Having constructed a basic model for the traffic inputs to the access network it is now possible to consider the way in which this model might be applied to analyse access network performance. In the typical AT&T ISTEL network access configuration, as illustrated in chapter 2, a number of users may gain access by dial up or direct connection.

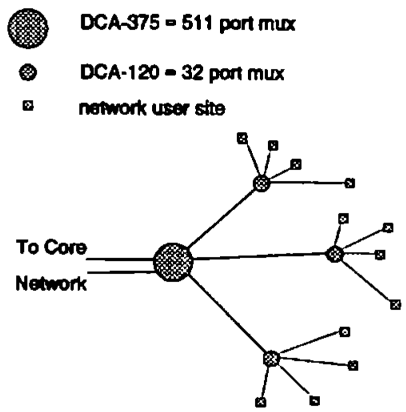


fig 3.6 The hierarchical connection of multiplexers

In order to model the characteristic delays to be expected from a given multiplexer configuration it is necessary to develop a process for combining the traffic statistics for the connected users. The combination is based on two features of the computer traffic, the message length and arrival probabilities. It is necessary to combine the message length and arrival rate probabilities for a number of different users in a way that produces a single model representative of the of traffic arriving at a multiplexer's inputs.

3.7.1. Composite Traffic Arrival Rates

In order to find the total composite arrival rate for a number, N, independent arrival rate processes it is necessary to find the total average arrival rate λ resulting from arrivals λ_1 λ_2 λ_3 .. . λ_N. The sum of N Poisson distributions produces a new Poisson distribution whose mean is equal to the sum of the individual means [Ozekici, 1990]

$$\lambda = \lambda_1 + \lambda_2 + \lambda_3 \ \dots + \lambda_N$$

Where a large number of users accessing a multiplexer can be classified into a small number of market sectors, the traffic profiles can be used in the summation process to quickly find the aggregate expected mean arrival rate. If a multiplexer had 3 Motor Industry users, 2 Leisure users and 1 Finance user then the total traffic arrival rate would be

$$\lambda_{Total} = 3.\lambda_M + 2.\lambda_L + \lambda_F$$

where λ_M = Mean Motor user arrival rate

λ_L = Mean Leisure User arrival rate

λ_F = Mean Finance user arrival rate

4.7.2. Message length Combination

In order to combine the message length distributions for a number of users to represent the overall aggregate message length distribution it is necessary to consider the way the distribution is represented. All possible message lengths are of integer value, and are represented by the reciprocal of the message service rate, μ, which is the time taken for the message to pass through the queue processor. As stated previously, due to memory limitations, the distributions are stored as a number of sampled points, with the probability of each of, say, 50 possible message lengths given. For N users there are distributions of message length $1/\mu_1$, $1/\mu_2$ ... $1/\mu_N$. In order to find the combined probability of each message length the weighted sum of the independent variables, message length, must be calculated. For the continuous case the following expression may be written for the average message length,

$$\frac{1}{\mu} = \frac{1}{N} [\frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{1}{\mu_3} + ...]$$

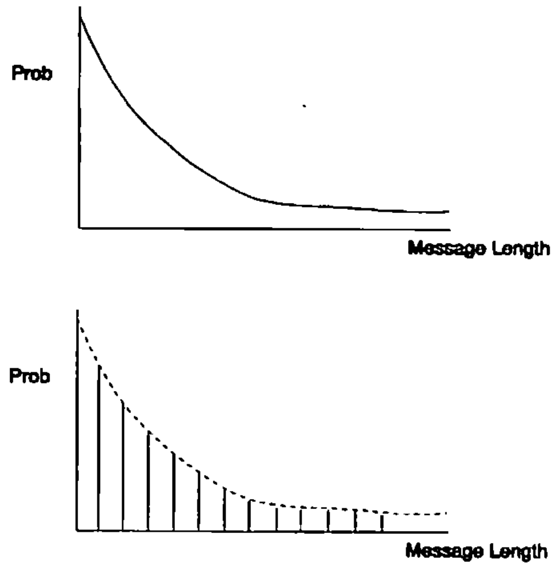


fig 3.7 The continuous distribution (above) is sampled (below)

For the sampled distribution this was best achieved on a piecewise basis, calculating the probability of each message length for every element, i, of the sampled distribution.
This may be written,

$$\frac{1}{\overline{\mu^{i}}} = \frac{1}{N}\left[\frac{1}{\mu_1^{i}} + \frac{1}{\mu_2^{i}} + \frac{1}{\mu_3^{i}} \cdots\right]$$

$$i = \{\ 1,2,3,4,5\ \ldots\ N\ \}$$

It is now possible to build up a single message length and arrival rate model based on the composite mean λ and $1/\mu$ values for the combined inputs of the multiplexer.

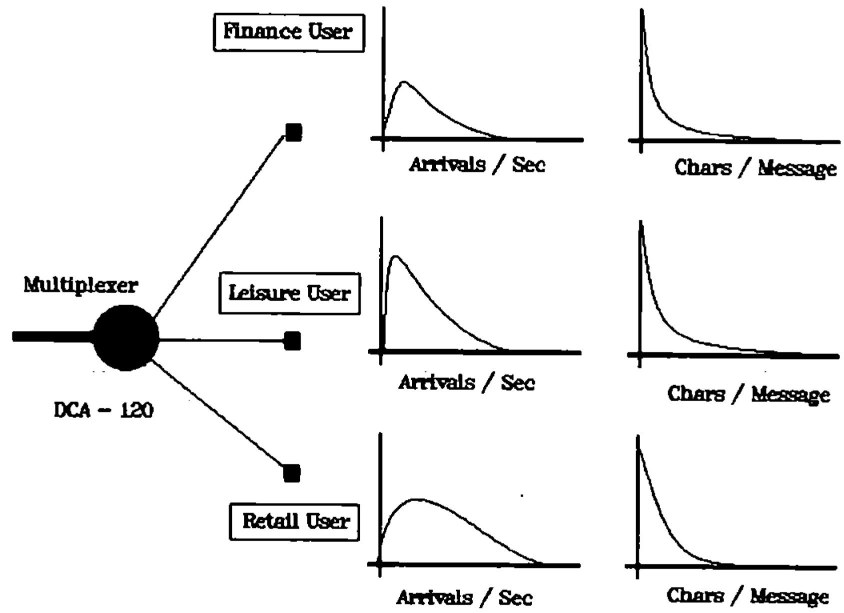


fig 3.8 Illustrating the combination of a number of user traffic profiles

3.8. Protocol Modelling

As indicated in chapter 2 the multiplexers (DCA-110, DCA-120 and DCA-375) use a data communications protocol to pass traffic between each other. The digital data communication message protocol DDCMP from the Digital Equipment Corp (DEC) is used between all DCA equipment.

The protocol increases the volume of data on the DCA links by adding a management data overhead. DDCMP operates on a frame basis between multiplexers and uses subframes, called 'plexs', to carry data for each terminal. The additional data added by DDCMP includes the following:

i. control bits for framing, delimiting the protocol packets;

ii. control bits for sub-framing, to delimit the data from each multiplexer input;

iii. addressing, to indicate the link from which data originates;

iv. error detection, to allow retransmission of corrupted data.

In discussing the modelling of the access network the focus is placed on the protocol between all DCA core multiplexers which is DDCMP.

3.8.1. The DCA Protocol - DDCMP

In order to evaluate the increase in line traffic due to the DCA inter-nodal protocol some simple equations are developed to estimate the increase, based on the rules of the protocol. First, it is necessary to examine the basis of DDCMP.

As previously mentioned when describing the sequence of events following the typing of a character, the protocol adds extra data bits to that character to indicate from which terminal and multiplexer the data originated. An important parameter of the DDCMP protocol is the frame formation time, or FFT. This is the time during which the protocol assembles frames from all the characters arriving at the multiplexer's inputs. If the frame becomes full before the frame formation time has elapsed then the packet is transmitted and the frame formation timer is reset. If a packet is not completely full and the FFT has elapsed then the system sends the packet regardless. A typical frame formation time is of the order of 200mSec. Even if no data arrives within the FFT an empty frame is transmitted to indicate to the management systems that the link remains functional.

3.8.1.1. Protocol Plexs

In the DDCMP protocol a plea has a maximum size of 7 characters, this includes a single character to denote the terminal from which the plea originates. The number of plexs created will be determined by how many characters arrive within the frame formation time and how many terminals are sending data.

4.8.1.2. Protocol Frames

In the DDCMP protocol the frame consists of:

i. a header, to indicate from which multiplexer the frame originated;

ii. all the plexs formed from characters transmitted from terminals within the FFT;

iii. a frame 'tail' of check bits to provide a measure of frame error detection.

The call destination is defined when the call is setup. This is performed by the user sending a message instructing the multiplexer to which destination node it wishes to 'connect'. The multiplexer has an internal microprocessor that establishes the call by testing to see whether connection is possible, i.e. whether there is a virtual circuit available to the destination and that it is not busy.

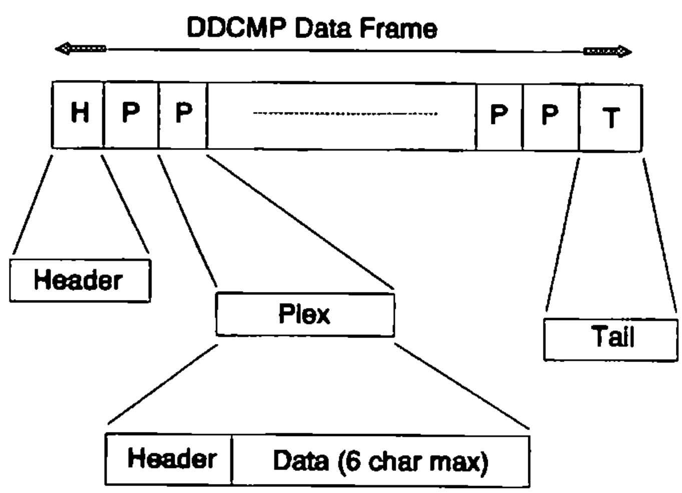


fig 3.9 The format of a DDCMP protocol frame

3.8.2. Statistical Representation of the DDCMP Protocol

In order to account for the increase in DCA network traffic due to the DDCMP protocol some equations have been derived that can be applied to the traffic figures to adjust the message length and arrival rate probability distributions accordingly.

The total additional protocol characters added to the original user traffic profile is based on the actual values of frame (message) length and arrival rate over the range of values they take. Some simple integer arithmetic is performed to calculate how many data overhead characters would be produced for each possible arriving message length for all the users connected to the multiplexer. The basic model is then adjusted to account for this increase in traffic.

Referring to fig 3.9 for the DDCMP protocol:

h = Plex header length;

P = Plex length;

F = Frame envelope length;

H = Frame Header Length.

Define a function to account for the additional plexs or frames that are only partially filled, relating it to the result of the 'modulo' (MOD) function:

K is a variable that can only take on the value 0 or 1;

K = 0, if the MOD function has a non-zero remainder;

K = 1, if the MOD function has a zero remainder.

In order to write an expression defining the total number of characters leaving the multiplexer output, the basic input arrival rates must be known and the number of additional characters created by the protocol must be calculated. To simplify the analysis, consider only a single input, with arrival rate λ and message length σ_c characters, or σ bits. Where $\sigma = 1/\mu$ and μC represents the message service rate in message/bit.

The total overhead characters per second due to the protocol will be due to the creation of full and partial plexs and the resultant frames.

i. Define N_c to be the total number of characters, on average, offered by the multiplexer inputs for insertion into a protocol envelope per second.

N_c = total number of data characters + number of characters due to plex headers,

total data chars per second = arrival rate * message length = $\lambda.\sigma_c$

$$N_C = \sigma_c.\lambda + \lambda.\ [\ (\sigma_c \text{ DIV P}) + \text{K}\ ].\text{h}$$

The characters due to the plex headers are caused by each arrival being assigned to its own plex. Therefore a minimum of one plex per arriving message is required. The number of plexs per message is defined by the message length divided (using the integer divide function DIV) by the plex size + one extra plex to carry the remainder of characters not exactly filling a plex (hence the K term), if one exists.

The overhead due to the plex headers = $\lambda.[(\sigma_c \text{ DIV } P) + K].h$

ii. Define the total number of protocol frames per second to be N_e,

N_e = total number of characters (including plex headers) divided by the frame size, the K function means that the partially filled frame is accounted for.

$$N_e = [N_c \text{ DIV } (F\text{-}H)] + K$$

iii. Let Δ be the total character overhead due to the protocol,

$$\Delta = \lambda.[(\sigma \text{ DIV } P) + K].h + Ne.H$$

The packet formation time is typically less than one second and the overhead calculation must be modified. The overhead is based on the 'characters per frame formation time' basis and is then multiplied by the number of frames per second to give protocol overhead per second. This modified traffic figure is used to calculate multiplexer queuing delays.

3.8.3. Modelling Traffic Time Delay

For a queue with Markovian arrivals (Poisson distribution) and message lengths (negative exponential) with a single queue server, M/M/1 (Kendal notation), the mean service time + queuing delay, E[T], is given by [Schwartz 1977];

$$E[T] = \frac{1}{\mu.C - \lambda}$$

where the output channel transmission capacity is C bits/sec.

In order to estimate the delay for the range of traffic arrival rates and message lengths from the sampled distributions an assumption must be made. In essence it is necessary to calculate the expected delay for each and every combination of λ and $1/\mu$ from the sampled distributions. This effectively means that for all possible input traffic levels the queuing delay is calculated. In order to do this it is necessary to assume that each value from the sampled distribution may be used independently. This requires that the samples themselves are considered to be the mean values of separate distributions. It is accepted that the use of the assumption may introduce some error but the form of this analysis and the results it produces are significantly more useful than a single mean value result. It is important to gain some knowledge of the delay performance of the multiplexer and a measure of the relative values of delay that are sufficiently useful to justify any compromise in accuracy.

In order for this work to find the expected delay for a sampled distribution of range of arrival rates and message lengths, this approximation is deemed acceptable.

The delay calculation is therefore calculated on an elemental basis for the range of all λ and $1/\mu$. This will naturally lead to a three dimensional result with axis in delay, λ and $1/\mu$. The delay may therefore be plotted as a surface that describes the delay expected for each possible combination of message length and arrival rate. The surface can be considered to be the locus of the operating point for the network traffic, i.e. as the network traffic load changes, the arrival rate and message lengths are continuously changing and the surface describes the expected delay for the instantaneous operating point. This is not an entirely accurate picture of the system, because at time t_n the locus does not take account of the queue's state at time $t_{(n-1)}$. Again, this may be considered to be an acceptable inaccuracy because the manner of the analysis and the range of the results provided are useful.

3.9. Explanation & Interpretation of results

The NETAL package was written to support the work of this thesis. It provides a means of storing traffic arrival rate and message length profiles on disk and allows the user to create models of multiplexer inputs based on mixtures of network customers from differing market sectors. The multiplexer protocol and aggregated traffic profiles are used to estimate likely queuing delays.

The NETAL package calculates and displays the expected delay figures as a three-dimensional (perspective) plot, the x axis represents the message length and the z axis the arrival rate. The y plane shows the expected delay. In order to illustrate the interpretation of these graphs an introduction to their understanding is given.

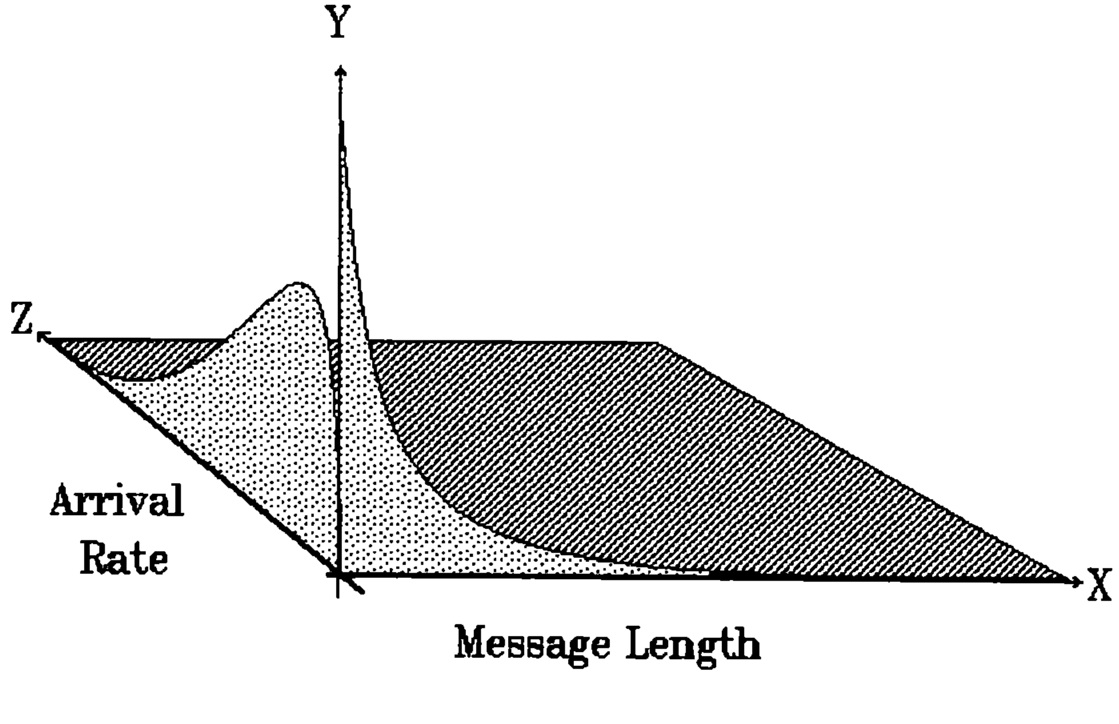


fig 3.10 Graphical representation of a two variable function

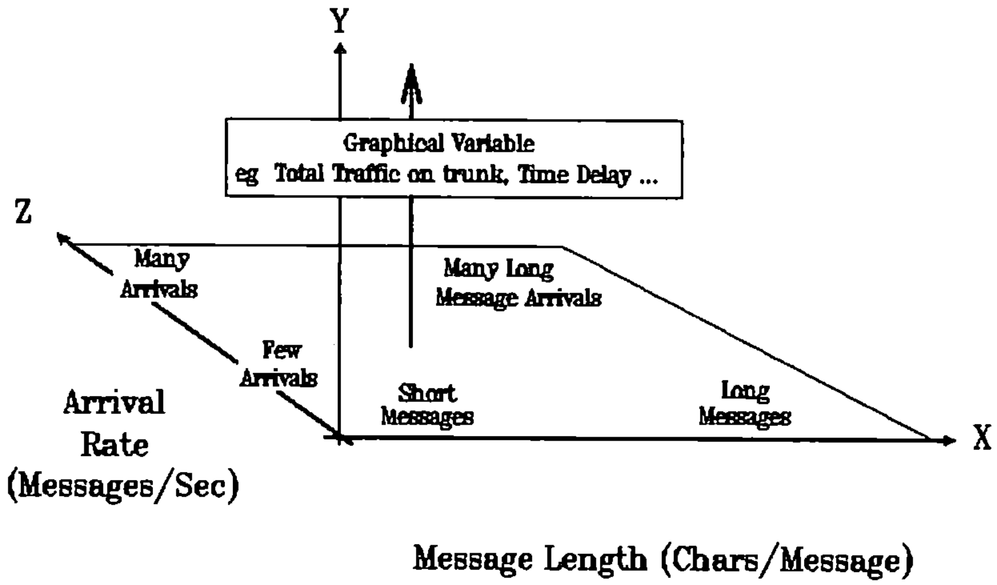


fig 3.11 Clarification of the message length, arrival rate, time delay axis

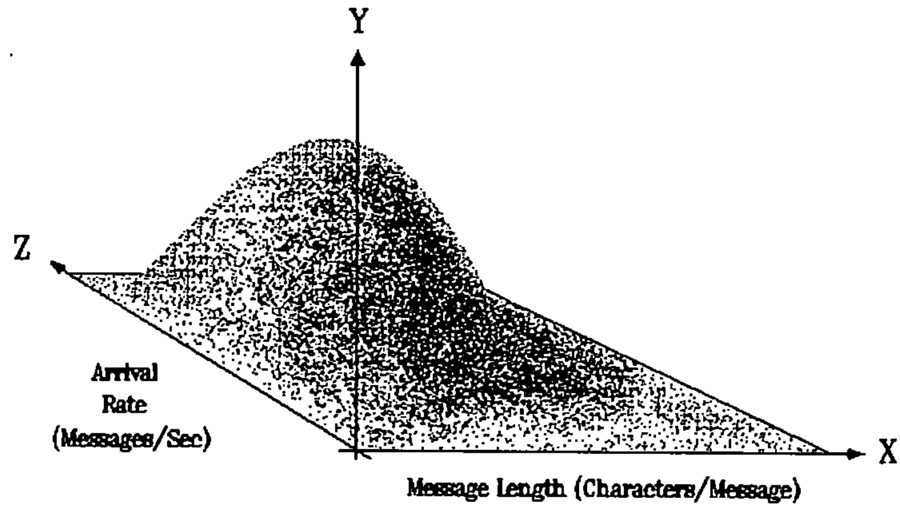


fig 3.12 Illustrating a smooth transition from one delay figure to another for varying traffic levels

i. If the delay surface appears smooth one may deduce that, as the traffic levels change, the operating point moves smoothly from one delay value to another. The access network may therefore be said to be working 'smoothly', there are no rapid, large changes in time delay for characters passing through the multiplexer. Figure 3.16 shows an example of this type of analysis. 12

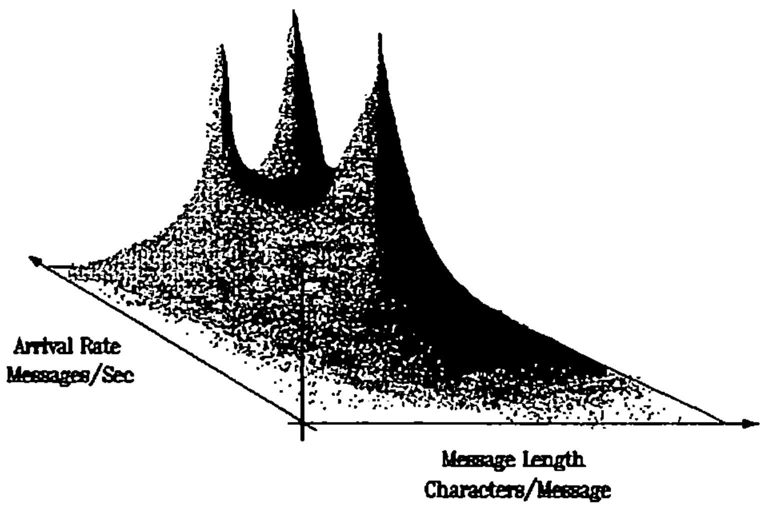


fig 3.13 Illustrating a rapid change in delay for small changes in traffic levels

ii. A heavily peaked delay surface represents the rapid transition of time delay from one value to another as the traffic levels fluctuate by small amounts. The network would suffer large changes in delay and would appear congested to the user as some traffic takes substantially more time to return from the host than others. Figures 3.17 and 3.18 show examples of this.

An important psychological effect of large random fluctuations in network delay is that the user becomes frustrated with the long delays because they compare so poorly with those traffic bursts that suffer little delay. It was found from surveying network users and by experimentation that a large but constant delay is subjectively preferable to a lower mean delay with a large variance.

3.10. Probability of Traffic Dependent Time Delay

The three dimensional graphs presented thus far show plots of expected queuing delay for a given range of arrival rate and message length. What is required by the network designer is a less complex representation of the delay performance for any given multiplexer configuration. in order to estimate of the percentage probability of the delay being less than a pre-defined threshold the probability of each possible arrival rate and message length combination occurring are required.

3.10.1. The Joint Distribution Function

The view of the 'delay surface' must be tempered with the realisation that the delays represented by each point on the graph are not equiprobable. That is, the probability of each combination of message length and arrival rate will dictate the probability of the traffic operating point being at each point on the surface. This combined probability of each arrival rate and message length is defined by the Joint Distribution Function, JDF. A typical JDF calculated by the NETAL package is shown in figure 3.19.

If the JDF of the total input traffic is known and the delay figures as a result of each traffic level are also known then it is possible to calculate the total probability of the delay being within specified ranges. Since the delay is defined on the y axis the surface may be intersected at any point in the xz plane and the area of the intersection will indicate all values of arrival rate and message length which lead to a delay above the threshold. These values of message length and arrival rate can then be tested against the JDF to reveal their probability of occurrence.

In order to visualise the problem, if the network designer wishes to know the probability of the delay exceeding a threshold then he may consider a horizontal slice being made through the delay graph at the level of that delay. The section of the graph above the cut represents the delays exceeding the threshold and that below it where the delay is better than the threshold.

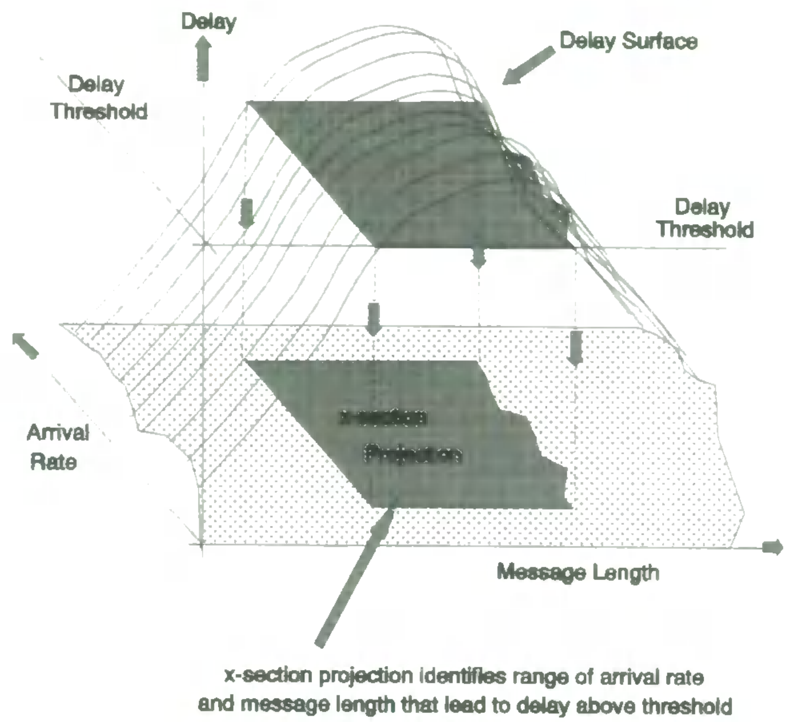


fig 3.14 Slicing the delay surface to find probability of threshold being exceeded.

The values of message length and arrival rate, MLAR, for which the delay threshold are exceeded is projected onto the zero delay axis. Mathematically this is performed with the Kroenecker function, which is used as a simple switching decision function taking on only two integer values either 0 or 1.

$$\Gamma(\lambda, 1/\mu) \quad = 0 \text{ for delay} > \text{threshold}$$

$$= 1 \text{ for delay} <= \text{threshold}$$

The salient feature of the JDF is that it describes the entire probability range of <u>all</u> (100%) occurrences of traffic loading, i.e. the volume under the surface must sum to unity. This is an important point to make because it is the central argument in estimating delay probabilities using this technique.

In order to calculate the percentage of the time that the time delay exceeds a certain value, the Kroenecker function is now applied to the JDF. The joint probability of the MLAR traffic values are summed for all '1' entries in the Γ function. If every delay was greater than the threshold then every value of Γ would be '1' and the probability would sum to unity, i.e. 100% probability. Consequently, for only the values of delay that exceed the threshold, where Γ = 1, the JDF values are summed to yield the total probability. It is also important to remember that both the message length and arrival rate probabilities tend to zero and the extremes of the area under the surface. Since the JDF is also the product of the two, the importance of the extremities is very small and they may be safely neglected without leading to any significant distortion of the probability figures. For example, if the probability of message lengths greater than 1000 characters is of the order of 1:10,000 and the probability of their arrival rate being less than 0.1 per second is of the order of 1:100,000. Though this arrival rate and message length might lead to a delay of 14 seconds (assuming 7 bits/char and 1200bps line speed) the probability of its occurrence is $1{:}10^{-9}$.

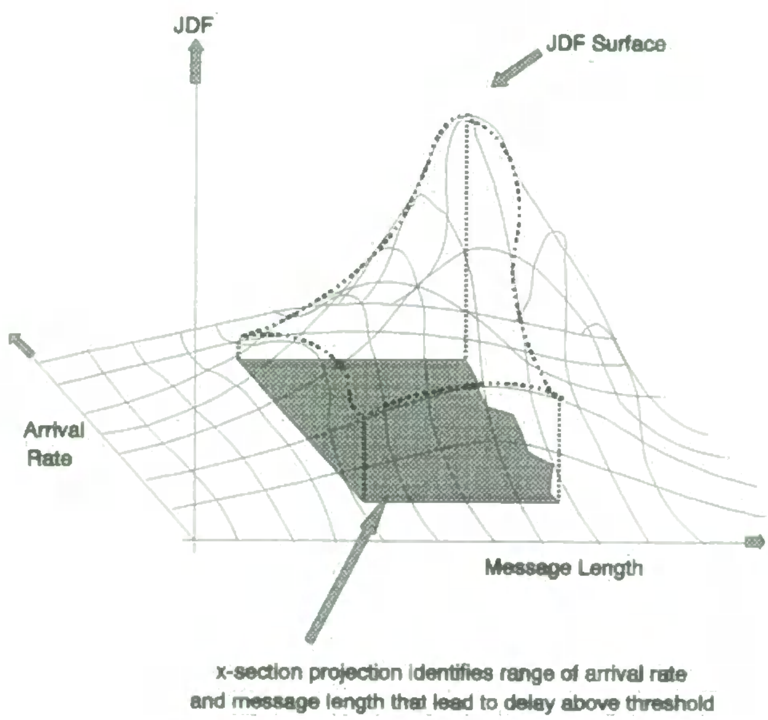


fig 3.15 Project arrival rate & message lengths leading to delay above threshold onto the JDF

The outline method is as follows:

i. On the delay surface, note all the combinations of arrival rate & message length that lead to a delay exceeding T milliseconds;

ii. sum all JDF probabilities for points where the corresponding delay exceeds T.

Define a binary function $\Gamma(\lambda, 1/\mu) = 1$ for $D(x,y) > T$,

0 otherwise;

iii. calculate the percentage probability of delay > T .

i.e. the total volume under JDF found in step ii;

iv. repeat from step i for a range of time delay thresholds, $T_1..T_{max}$.

A two axis graph can now be plotted showing delay thresholds, T, against the probability of the traffic delay exceeding each threshold. Figures 3.19 and 3.20 show examples of this. The NETAL package has produced many such graphs and these two examples clearly show that the percentage based delay figures gives a good indication of the access network performance. Figure 3.20 shows that, for approximately 80% of the time, the delay is less than 282mS yet for 95% of the time it is less than 300mS. i.e. the variability of the delay is not large. In contrast, figure 3.21 shows a delay of better than 791mS for 79% of the time but a delay of better than 1230mS for about 95% of the time. The dramatic swing in delay within the same percentage bound would indicate a much more variable delay performance and shows that the worst case figures are likely to be unacceptable to the user.

3.11. Results of the Delay Evaluation

The surface of the graph represents the time delay for the corresponding data message length and arrival rate (MLAR). Over a period of time, for any combination of MLAR the delay will follow the surface, which therefore represents the locus of the traffic delay operating point.

The analysis results are shown in the following figures 3.16 to 3.21.

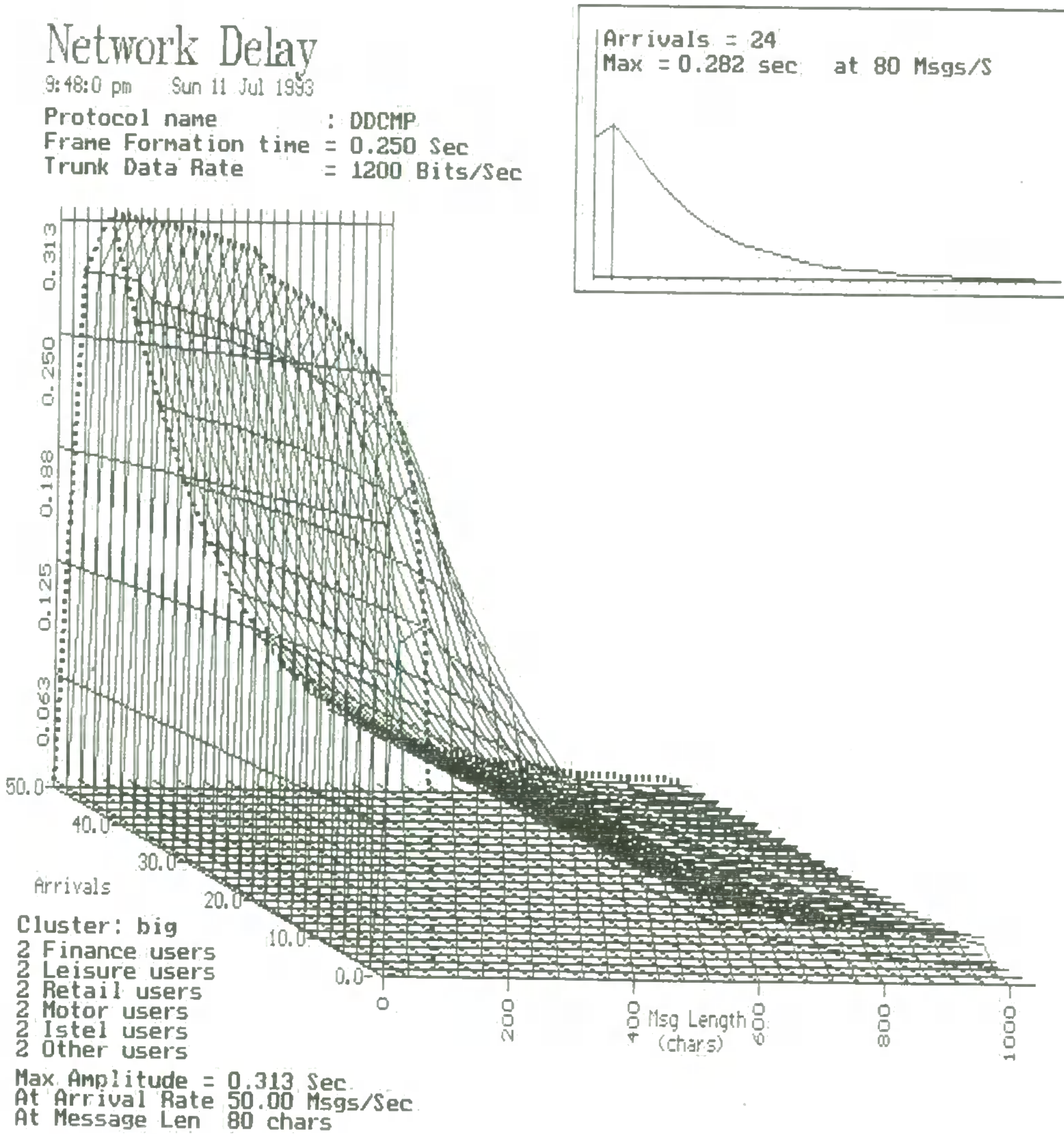


fig 3.16 The typical data delay as a function of arrival rate and message length

Typical access multiplexer delays figures using DDCMP protocol values, the data rate is 1200 bits/sec and the protocol forward time is 250 milliseconds. Note that the surface is smooth, i.e. there are no rapid changes in delay for small changes in arrival rate or message length.

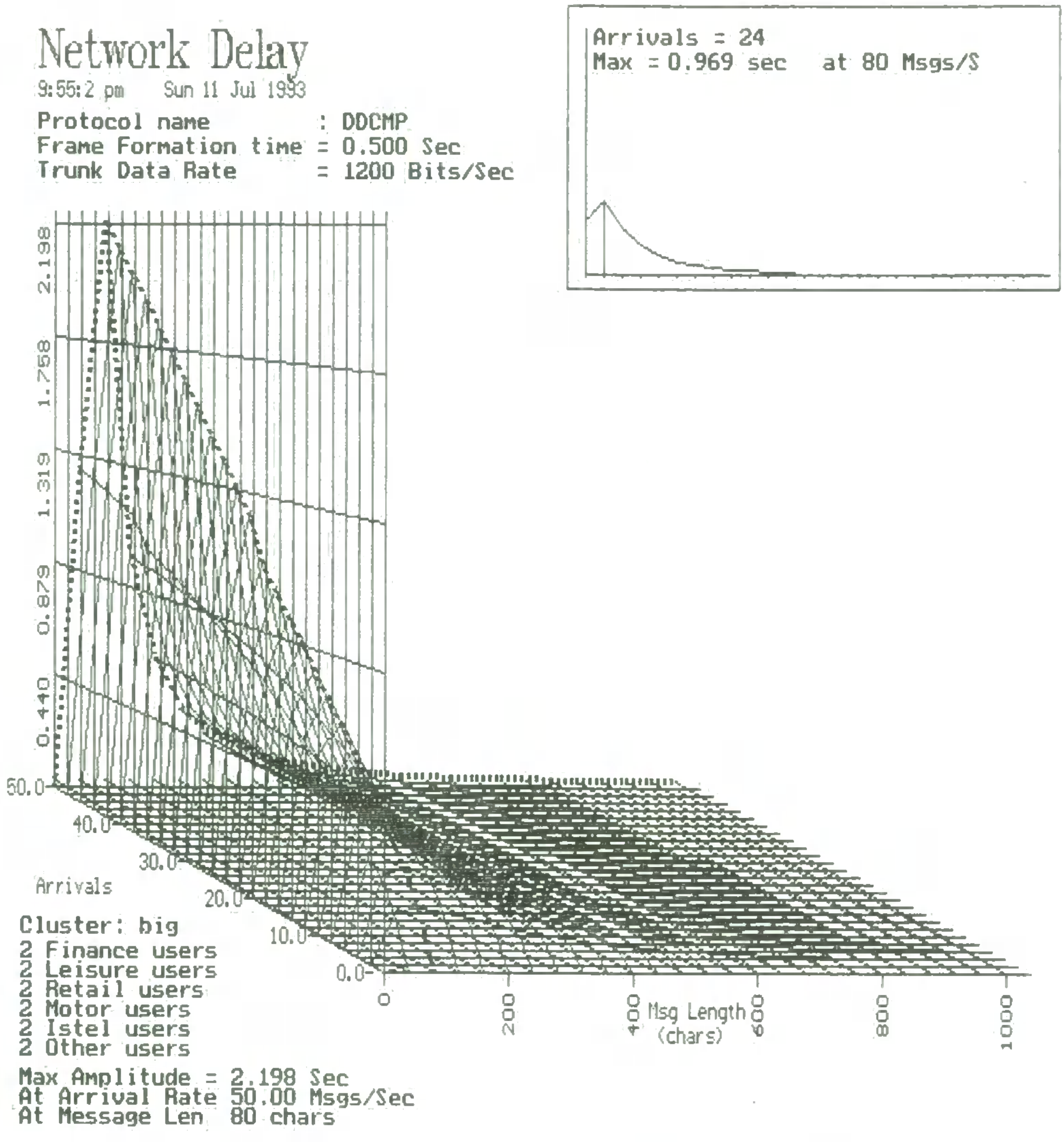


fig 3.17 Illustrating the effect of increasing the protocol forward time

Note that the graph slope becomes more steep and the delay for high arrival rates increases.

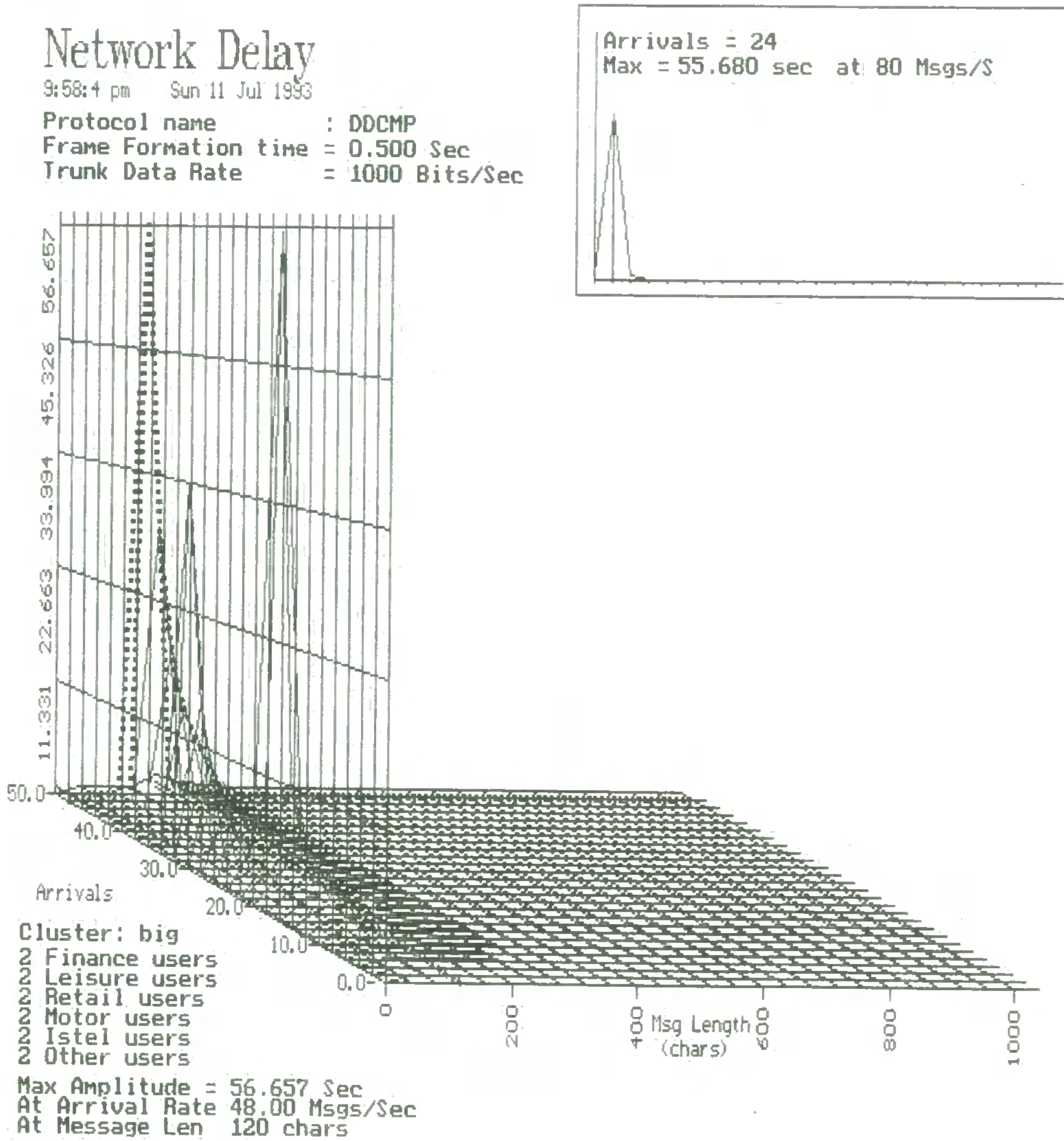


fig 3.18 The effect of lowering the data rate and retaining a large packet forward time

Note that the graph becomes very peaky, the maximum delay rises to 56 seconds, this is characteristic of a congested network, where the data transmission capacity is insufficient.

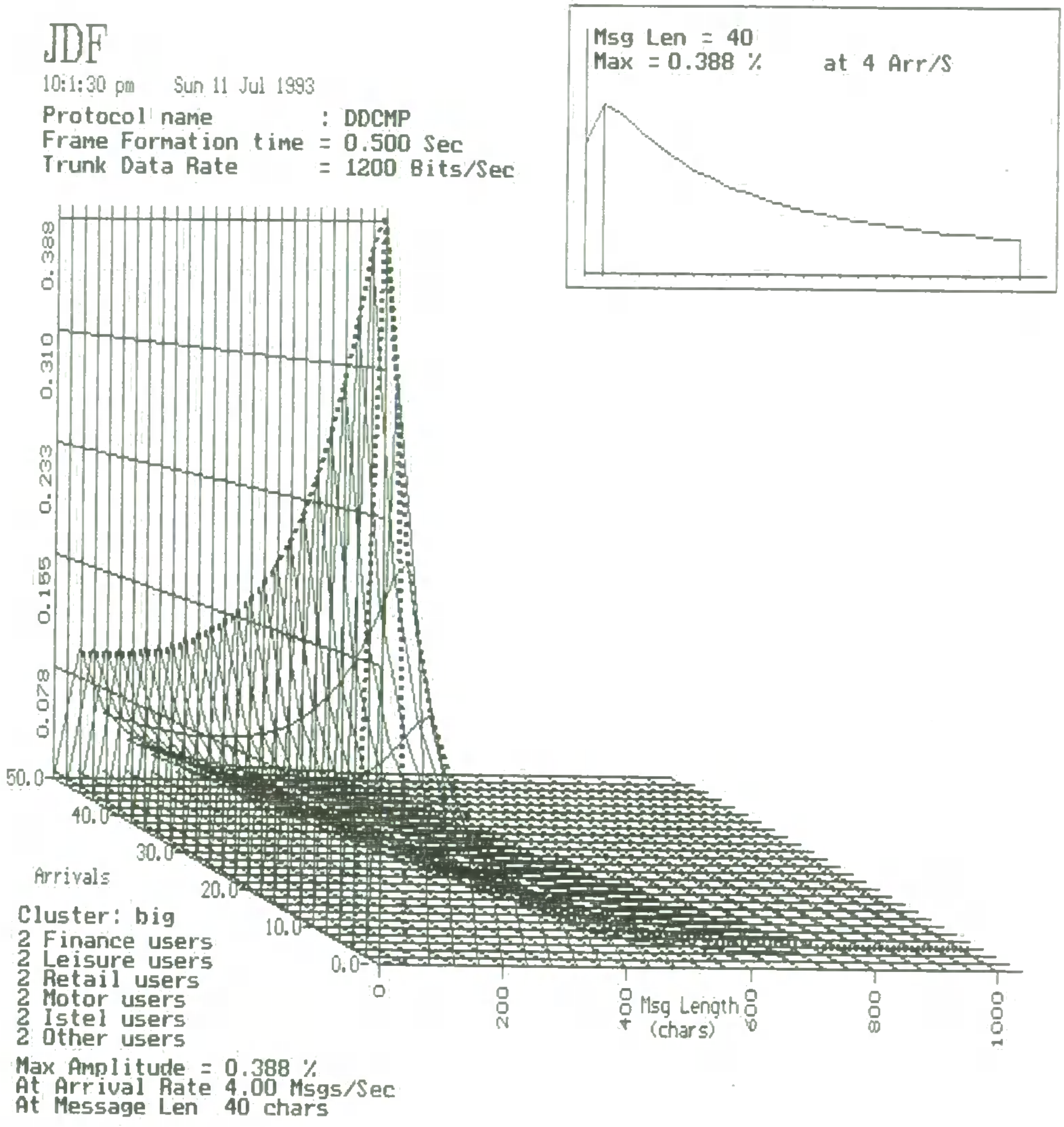
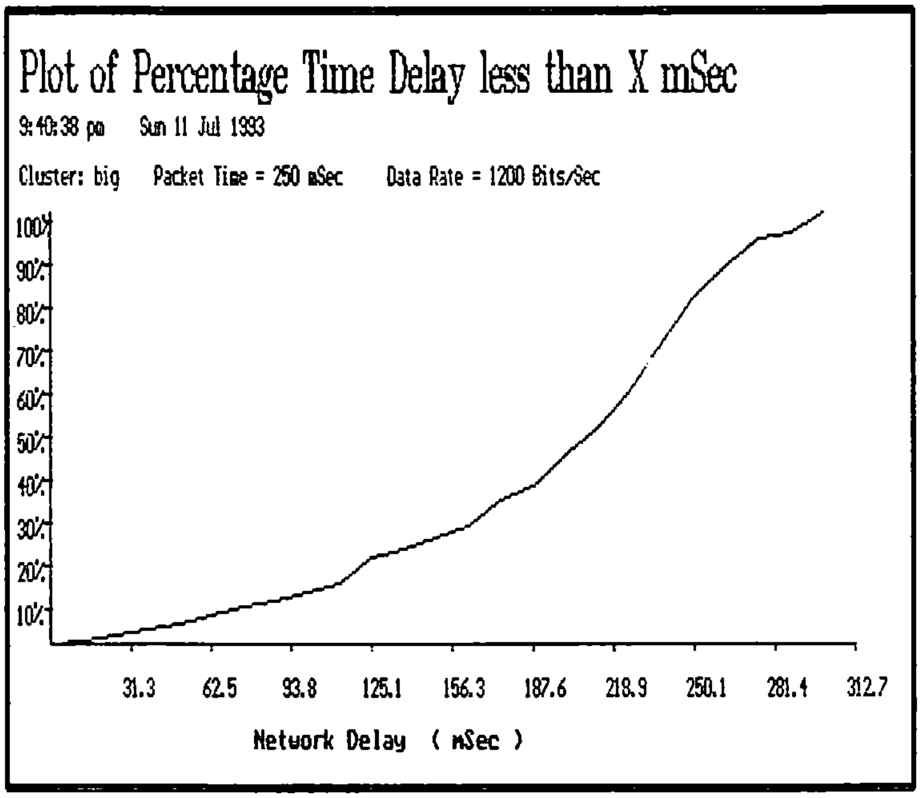
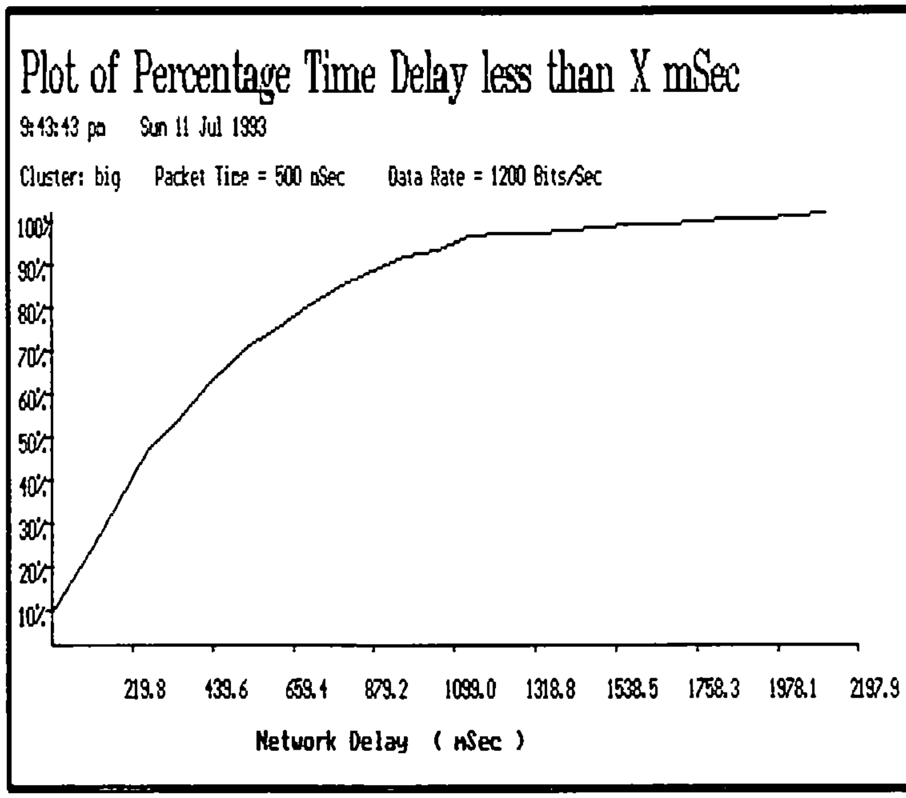


fig 3.19 The typical joint distribution function from the NETAL package



Delay < 12.5 mSec for 0.4% of time Delay < 25.0 mSec for 1.3% of time
Delay < 37.5 mSec for 2.4% of time Delay < 50.0 mSec for 3.9% of time
Delay < 62.5 mSec for 5.1% of time Delay < 75.0 mSec for 6.8% of time
Delay < 87.5 mSec for 9.0% of time Delay < 100.1 mSec for 10.5% of time
Delay < 112.6 mSec for 12.4% of time Delay < 125.1 mSec for 14.2% of time
Delay < 137.6 mSec for 20.5% of time Delay < 150.1 mSec for 22.3% of time
Delay < 162.6 mSec for 25.1% of time Delay < 175.1 mSec for 27.5% of time
Delay < 187.6 mSec for 33.3% of time Delay < 200.1 mSec for 36.7% of time
Delay < 212.6 mSec for 44.6% of time Delay < 225.1 mSec for 50.5% of time
Delay < 237.6 mSec for 58.9% of time Delay < 250.1 mSec for 70.3% of time
Delay < 262.6 mSec for 80.7% of time Delay < 275.1 mSec for 88.2% of time
Delay < 287.6 mSec for 94.2% of time Delay < 300.2 mSec for 95.4% of time
Delay < 312.7 mSec for 100.0% of time

fig 3.20 Percentage time delay graph for the delay surface in figure 3.16



Delay < 87.9 mSec for 7.4% of time Delay < 175.8 mSec for 19.8% of time
Delay < 263.8 mSec for 32.1% of time Delay < 351.7 mSec for 45.5% of time
Delay < 439.6 mSec for 52.8% of time Delay < 527.5 mSec for 61.7% of time
Delay < 615.4 mSec for 68.9% of time Delay < 703.3 mSec for 74.0% of time
Delay < 791.3 mSec for 78.7% of time Delay < 879.2 mSec for 84.0% of time
Delay < 967.1 mSec for 87.2% of time Delay < 1055.0 mSec for 90.3% of time
Delay < 1142.9 mSec for 92.0% of time Delay < 1230.8 mSec for 94.8% of time
Delay < 1318.8 mSec for 95.4% of time Delay < 1406.7 mSec for 96.0% of time
Delay < 1494.6 mSec for 96.5% of time Delay < 1582.5 mSec for 97.0% of time
Delay < 1670.4 mSec for 97.5% of time Delay < 1758.3 mSec for 98.0% of time
Delay < 1846.3 mSec for 98.4% of time Delay < 1934.2 mSec for 98.8% of time
Delay < 2022.1 mSec for 99.2% of time Delay < 2110.0 mSec for 99.6% of time
Delay < 2197.9 mSec for 100.0% of time

fig 3.21 Percentage time delay graph for the delay surface shown in figure 3.17.

3.12. Application of the NETAL Results

The NETAL package is useful in the network planning department when it is necessary to determine to which multiplexer a new customer should be connected. The expected delay figures can indicate whether a proposed mixture of customer profiles on any given multiplexer is likely to ensure acceptable delay performance, based on either 80^{th} or 90^{th} percentile figures. It is also useful for estimating how many additional users may be added to a multiplexer that is lightly loaded. For instance the NETAL package might indicate that delays are within a 90^{th} percentile of 700mS if twelve Leisure users are added, or six Motor users or twenty Finance users. This allows network planners to determine new connections and also impending needs for new multiplexers.

CHAPTER 4

Core Network Analysis and Modelling

4.0. Introduction

This chapter analyzes core networks, firstly from the point of view of human perception and then follows with an examination of the component features such as rings and meshes. It considers the fundamental properties of networks which are analyzed in order to determine what constitutes good network design practice. A review of methods employed by other researchers serves to form the basis of the development of ideas in the creation of a new optimisation system described in later chapters.

This design programme aims to determine the optimal capacity trunks between nodes, the selection of the nodes themselves is assumed to be known in advance. In provisioning an international large scale network the issues in node selection are determined by market analysis and corporate objectives, and are therefore outside the domain of core network design. Further, the cost of nodes is not considered since the large capital cost is largely fixed for most likely configurations. The dominant costs are likely to be related to physical location, such as office rental and the deployment of operational and maintenance staff.

For the purpose of the following work, with reference to service availability as discussed in chapter 2, it is assumed that a single backup path is required for all traffic demands. Link disjoint paths, sharing no common links, are provided for primary and backup routes. This ensures resilience to any single link failure and the probability of additional simultaneous link failure is deemed to fall outside guaranteed service levels. The network technology is assumed to provide automatic circuit rerouting following the detection of link failure. This is a feature of the NET, IDNX, TDM system currently used by AT&T ISTEL throughout Europe. All traffic demands are specified in terms of permanent dedicated channels between end nodes, consistent with the NET TDM network technology.

4.1. Human Perceptions of Network Designs

Only a few network designs are published in the academic literature, the most notable being those of the ARPAnet (Internet) [Frank, Frisch & Chou 1970], [Frank & Chou 1974] which is a working testbed for the development of network techniques and technology. While the ARPAnet is a packet switched network it is instructive to analyse some of the features of the designs and their objectives. It is extremely difficult for the human observer to contemplate the quality of the designs presented due to the extraordinary number of variables and possible outcomes of the network design problem. Even when an analysis of the design results has been conducted the quality of the network topology is limited by the ability to analyse on the basis of human perception of performance and cost criteria. An inescapable fact of any analysis is the layer of human interpretation of the results that takes place and leads to further design perturbations and new design methods.

Hierarchical network designs have come about for two real reasons, firstly because they are potentially solvable, due to the partitioning of the large problem into smaller sub-problems, and secondly because human understanding of them is possible. It is instructive to consider the way in which human perception of network design functions. It is not apparent that a broad discussion of the basis of network design has been published before, the start of network design methods is usually 'let us minimise the general cost function stated below ...'. This is then followed by an attempt to reduce the mathematical complexity of the resultant ing equations which may potentially hide any inherent opportunities for reduction by a more clear understanding of the problem.

In searching for occasion to reduce the network design problem's complexity it is necessary to understand how the human's eye-brain combination has certain 'beliefs' and 'modes of contemplation'. These influence the perception of the problem but are not necessarily correct, it is possible to identify some of the features of good and possibly bad designs and predict where the human level of understanding is weak. This can be summarised broadly as follows.

4.1.1. The Eye has a Natural Tendency to Cluster Objects

The human eye-brain interaction is inherently very good at performing clustering. It is not known why there should be such an in-built biological mechanism, but it has most probably evolved as part of the 'survival mechanism'. The eye can visualise apparent density, where populations of objects (threats or food) are greatest and so the eye can then effectively see (or judge by experience) a centre of gravity. This may therefore correspond to a maximum probability of threat or food.

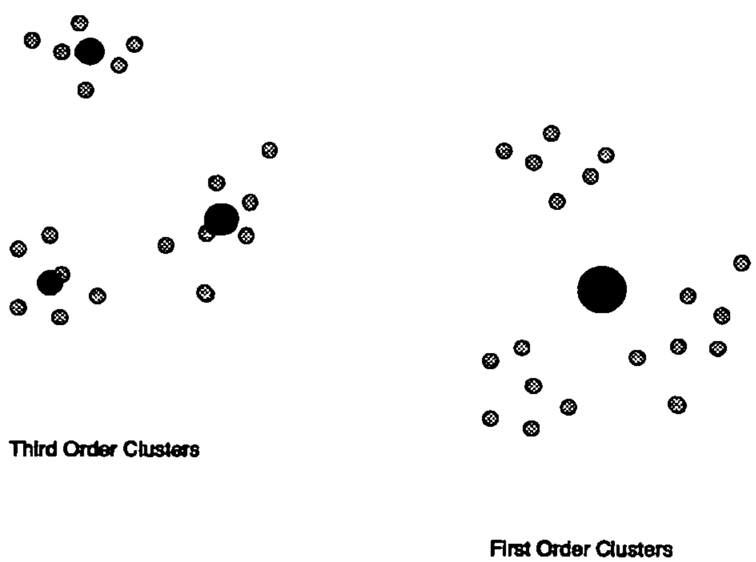


fig 4.1 Natural clustering by eye

The clustering effect influences a number of perceptions of network design. Primarily it appears that the human observer expects adjacent nodes to be connected together (a form of logical clustering), with the size of the connections (link capacity) being proportional to the number of nodes in the cluster.

The problem caused by this perception is that capacity and link requirements between nodes are entirely misconceived. The reasons are two-fold, firstly, the size of a traffic demand does not necessarily determine, merely influence, the link requirements. It is the source and destination, coupled with the size, of a demand that is vital to the specification of the internodal links. Secondly, the allocation of capacity to demands between nodes is usually accomplished by the use of routes with one or more links. The use of routes with more than one link greatly complicates the display of a network on a diagram and is therefore generally neglected. The eye has henceforth lost a vital clue to the correct interpretation of the network design.

4.1.2. The Eye Cannot 'See' Traffic on a Diagram

Perhaps the greatest problem with attempts to represent networks on a diagram, to aid human perception, is that it is very difficult to represent both traffic demand and traffic allocation. The importance of diagrams stems from the need for the human to 'understand' tables of figures. Perhaps the assumption that the diagram aids understanding is flawed and new ways of interpreting network configuration data are required. There is no doubt that very large networks are extremely difficult to comprehend, so that it is only by breaking down the 'views' of the topology into hierarchies or smaller regions that some semblance of understanding may be formed.

The eye will wish to interpret a traffic demand across a geographical area in the form of density of demand. Unfortunately the human perception of centres of density (or gravity) is based on real world experience where the vast majority of objects are uniformly dense. In the networking sense, when viewing graphical representations of node locations and traffic demands the perception, by eye, of a centre of gravity is poor. The eye is lead to expect allocation of links in the vicinity of demand and is unable to see the effect of applying demand to links along a route between each end-node. The eye's approximation of weighting to regions of heavy demand is therefore a dangerous instinct when the human appraisal of network design is attempted. The visual centre of gravity is unlikely to coincide with the traffic based centre of gravity or indeed the channelling of the data between nodes. The allocation of links between nodes is unlikely to correspond with the simplistic manner in which the eye views links.

It is possible to draw link widths on a diagram to try and convey an impression of loading, but this the fails to indicate the sizes of demands at the ends of links. The result of this is that the human designer will find it very difficult to evaluate the allocation of capacity on each link for all the required routes and any instant reactions upon seeing a machine generated topology should be tempered with this understanding.

The diagram below shows how it becomes difficult to visualise centres of gravity when points are weighted due to the increased complexity of balancing distance and weight cost for each point relative to all others. In effect the weights overrule the spatial relationships of the points to be clustered and the eye becomes confused.

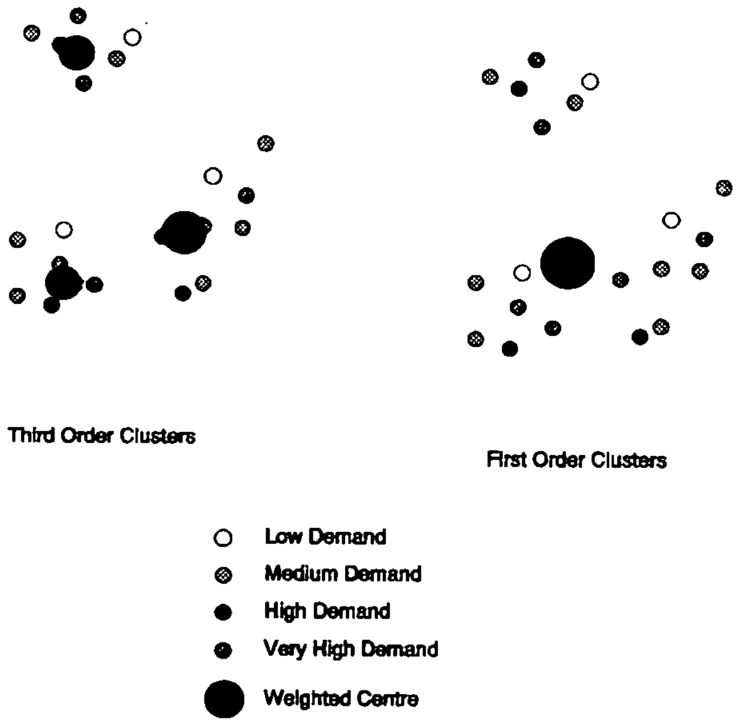


fig 4.2 Weighted clustering is difficult by eye

4.1.3. The Eye Wishes to See the Minimum of Links

A further human instinct that is applied when looking at networks is that of 'less is best'. Inherently, given two networks - one with more links than the other - the one with less connections will be seen as simpler, therefore less costly and hence more desirable. The human thinking tends to lead one to perceive less connections to be cheaper, yet many instances arose during the research period where plausible network designs were generated with more links of a complex nature but lower overall cost than those with fewer links. Another problem with a lightly connected network is that, by definition, the fewer links the more likely any one route is to use each link. The ring network characteristics prevail, therefore routes are longer, capacities are necessarily larger and, unless path lengths are minimised, costs are potentially higher. Though in any one instance it is the link cost structure that determines whether costs are higher, it is necessary for link setup costs to heavily dominate path length costs to reverse this situation. It is therefore clear that the link cost and traffic requirements are critical in influencing the type of network topology and each of the examples given may occur under particular circumstances.

The requirement for link disjoint paths may contribute to the proliferation of links in the topology if there is a heavy restriction on the number of hops per path. This is a result of it being more difficult to take advantage of high hop count paths and share capacity on trunks common to many other routes.

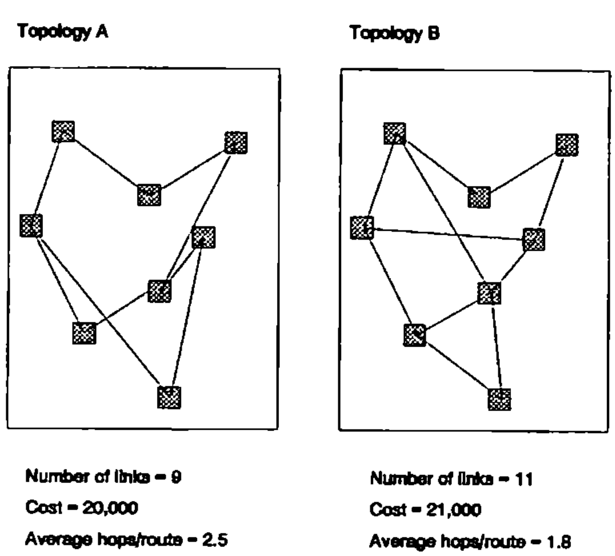


fig 4.3 Minimum links are not necessarily the optimal solution

The figure above shows how two topologies can have significantly different costs and characteristics the merit of each not being immediately obvious. Topology A might initially be selected on the basis of its lower cost. However, to illustrate the point of revenue potential, consider the number of links in topology B, the cost per link and the cost per likely route. Topology B is only 5% more expensive than topology A, for any network provider to make a selection between two topologies, a 5% cost differential should not be the one and only criteria. There are 11 links in topology B compared to 9 in A. The cost per link is lower and, perhaps most importantly, the average number of hops per path between all nodes is lower. This will mean that for the majority of cases where traffic demands between all nodes are of the same order topology B will offer a considerably greater revenue potential. It can carry more traffic. There would have to be a very special case for selecting topology A since there are very few nodes with a small number of hops between them for primary and backup paths. Topology A would be most suited to a low traffic volume packet switched network where it is essential to connect as many users together as possible without incurring large link costs.

4.1.4. The Eye does not Perceive Crossover Links to be Efficient

Following the previous understanding of fewer links erroneously appearing 'good' to the eye, the sight of cross-connected links are seen as highly unattractive. This is partially for reasons of apparent length, the cross-over links are likely to be at least half the distance across the network. A further reason for the eyes reluctance to accept cross network links appears to be the very fact that even a single cross-over looks somewhat complex.

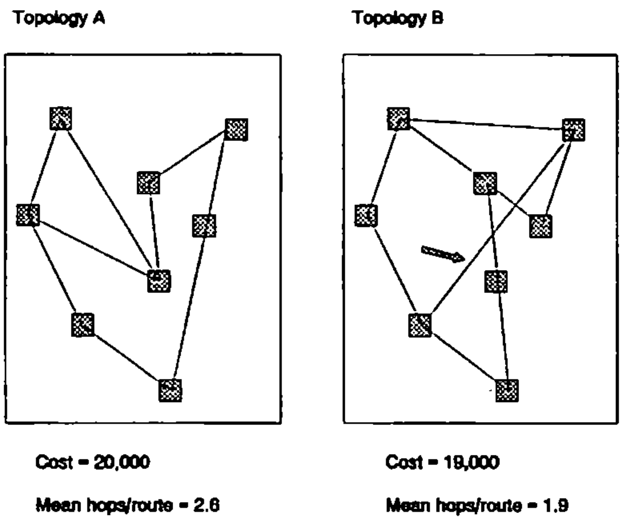


fig 4.4 Cross connections are advantageous

However, there are many advantages to having at least one cross network link. Cross links greatly reduce the number of hops required by routes to traverse from one side of the network to the other and they therefore have a positive effect on the inherent reliability of a network. For the same reason they also improve the revenue potential of the network. Another advantage of cross network hops is that in reducing the mean route length they reduce overall capacity required on links that would be congested with carrying traffic from long routes. The benefit of cross network links is also apparent if the load pattern should change. If a rapid growth in load should occur in a shorter time than it is possible to provision new capacity in one area of a network, the cross-network links can have the effect of dissipating the increased load and spreading it around the network within a small number of hops. It would not usually be possible to use loops in the same way since if any link within the loop is saturated then the path is effectively blocked.

4.1.5. The Eye Cannot Envisage Aggregation of Primary and Backup Routes

With large networks of more than a few nodes the network design problem becomes so difficult that it is not possible to visualise the way in which a large number of traffic demands lead to a multitude of primary and backup routes. It is therefore difficult for the brain to validate or even quantify the quality a network design presented to it.

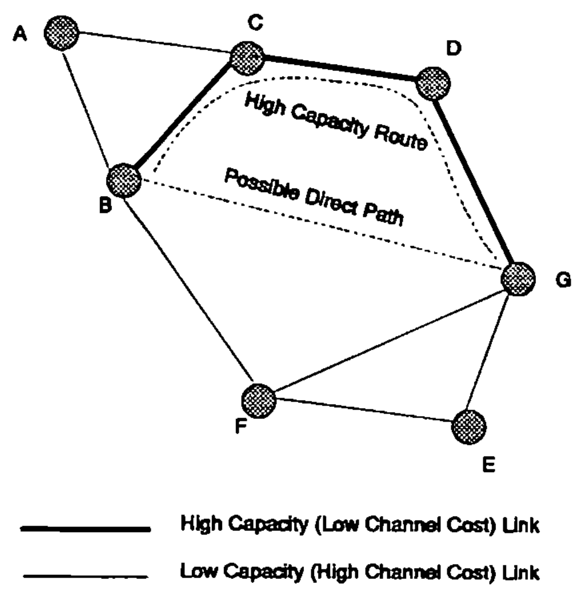


fig 4.5 High capacity links offer lower channel costs

From the network above it might be found that the eye is 'pleased' by the connections shown - the network appears reasonably efficient. The tariff must be such that the high capacity links offer sufficient channel cost savings to make up for the increased path length. In practice the tariffs are banded for ranges of capacity and there are significant savings in channel costs as trunk capacity rises. Inter-node distance alone therefore does not determine route selection.

4.1.6. Summary of Human Analysis

Having considered these perceptions it is possible to continue developing a design method, accepting that, while not every network design may be 'good', the human is ill qualified to pass judgement on the basis of a simple link diagram without the benefit of extensive numerical analysis. By necessity, for the benefit of rapid human comprehension, network designs are usually represented as diagrams showing links between nodes, with some means of denoting the link capacity. While these diagrams may be used as a guide to the nature of the network, even if a picture paints a thousand words, it may be insufficient.

4.2. Examination of Network Properties

There is important insight to be gained from examining the various forms a network might take to determine what topology features arise under different traffic and trunk tariff conditions. Particular interest is focused on the properties of the network topology that lead them to incur cost and carry traffic efficiently.

4.2.1. Ring-like Features in Core Networks

The pertinent feature of the ring network is that it represents the minimum number of links possible to join all nodes to all others with both primary and backup routes. For any non-trivial network the ring network directly translates to the Travelling Salesman Solution (TSP) network. The primary and backup paths meet the link-disjoint criteria.

There is no complex route selection required for primary and backup paths, except that the primary is the shortest distance around the ring from source to destination and therefore the backup takes the opposite route around the ring. No matter which route is considered, the overall mean route length will be M_R where;

$$M_R = \frac{(N-2)}{2}$$

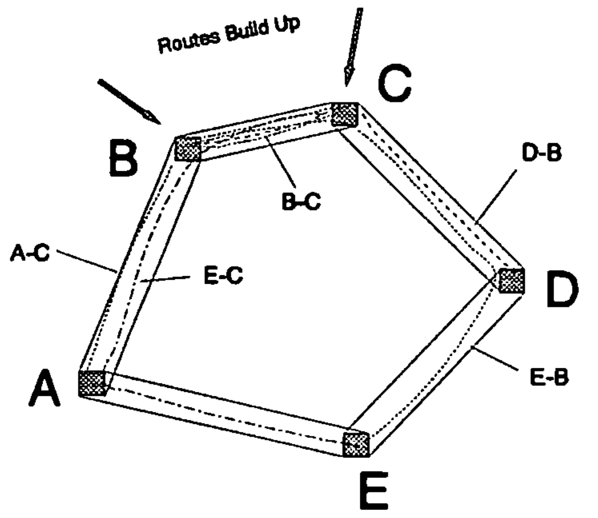


fig 4.6 Routes congregate on links in the ring

If the shortest path between each node pair around the ring is selected the build up of routes through each link can be seen to be dramatic and entirely due to the number of nodes in the ring. Looking at fig 4.6 it can be seen that the number of routes using the link indicated are entirely due to those terminating at the two highlighted nodes. The mean traffic level on links in a ring will therefore be much higher than for any other topology. This has an advantage in that the higher the number of channels required on each link the higher the capacity required and, considering the typical capacity/cost functions of fig 4.9 the lower the unit channel cost. The disadvantage of the ring is that the average number of hops per path increases with the number of nodes in the ring. As the ring increases in size the cost savings due to the channel cost reductions are lost through the increased path length around the ring compared with the direct path. The number of hops directly affects the service availability.

The mean traffic requirement of each link T_{ij} for the ring is,

$$\bar{T}_{ij} = \sum_{\substack{i=1 \\ i \neq j}}^{N} D_i + \sum_{\substack{j=1 \\ j \neq i}}^{N} D_j$$

where D_i is the demand to the i^{th} nodes;

assuming primary and backup paths are required.

This figure therefore represents the maximum link capacity required for N nodes in a ring network. The ring network offers the minimum trunk distance connectivity for N nodes but suffers the maximum number of hops possible for resilient connection. For a ring-like feature in a core network the positive cost benefit is the minimum trunk distance and the low hop count, the negative cost benefit is the high channel allocation and the increased path failure probability. The ring is useful in connecting a number of nodes together that have low traffic demands to the rest of the network, if trunk setup charges are high with respect to the distance related cost component then using the minimum number of trunks is cost effective when the setup charges of direct paths outweigh the rings increased total path length.

4.2.2. The Full Mesh Core Network

The most important feature of the mesh network is that all routes have a single hop primary path and a two hop backup path:

mean number of hops for all primary paths = 1;

mean number of hops for all backup paths = 2.

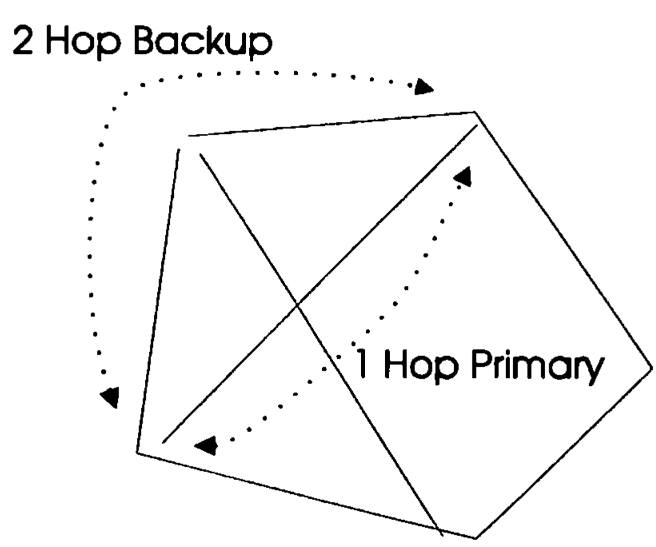


fig 4.7 1 hop primary and 2 hop backup of the mesh

Assuming primary and backup demands are allocated there will be 3 links required to satisfy each demand, one for the primary path and two for the backup path.

Therefore the mean traffic level on each link will be three times the mean of all demands:

$$T_{ij} = 3.\overline{D}_{ij}.$$

Since the fully connected mesh is the maximum resilience connection for N nodes this figure represents the minimum capacity requirement per link for N nodes under the constraints of providing link and node disjoint primary and backup paths. This figure represents the capacity requirement per link if backup capacity is allocated equal to the primary traffic.

The full mesh, in contrast to the ring, has a high total trunk distance and trunk count, this is therefore a significant cost factor. There is little increase in trunk capacity, relative to traffic requirements, to reduce unit channel costs as is possible with the ring. The advantage of the mesh topology is that each individual primary and backup path is of minimal distance, the trunk cost tariff will therefore determine whether there is a net benefit in mesh-like or ring-like topology. Over a large network it is likely that the traffic demands will work in concert with the trunk tariff to create loops in some regions while leading to full meshes in others and a hybrid in-between.

4.2.3. Revenue Potential of a Network

The notion of Revenue Potential of computer networks is proposed. This represents the amount of revenue a given network can generate for a given traffic demand. The Maximum Revenue Potential (MRP) of a network is realisable only when each traffic demand uses only one link and has no backup requirement. In this instance each link is entirely devoted to carrying each demand. The path length is a minimum, the unused capacity overhead is zero and hence the maximum fee paying traffic volume is carried.

Since the maximum revenue potential of a network is confined to an impractical scenario it is necessary to determine a way of calculating a realistic RP for a given network. A way of achieving a more realistic calculation of the network RP requires that primary and backup route hop mean values are known. This then allows the designer to estimate the number of customer channels that can be carried across the network with the known mean route hop figures since it is possible to divide the total network capacity C_{TOT} by the total of the mean primary and backup hops.

$$RP_{TOT} = c_u \cdot \frac{C_{TOT}}{(\ m(P) + m(B)\ )}$$

where RP_{TOT} = total revenue potential;

c_u = unit cost per path

m(P) = mean channel usage per primary path;

m(B) = mean channel usage per backup path.

The equation above uses the primary and backup channel usage as a measure of the network resource utilisation by allocated paths. There is a subtle distinction between channel usage and hops since the work of Agarwal [1989], who shows the backup channel requirements are not the same as the primary channel requirements. This subject is dealt with in more detail in chapter 6.

The use of this revenue potential figure places the onus on the network provider to gather figures for the mean number of channels or hops per primary and backup route from an analysis of past topology and customer traffic requirements. This is entirely practical if the provider has been running a growing network for a significant period of time and has a well developed management system.

By way of example, consider a notional network with:

i. 10 nodes;

ii. 16 links of 32 channel capacity;

iii. 8 links of 8 channel capacity;

iv. the provider has a history of providing 2.0 hops for primary routes and 2.5 hops for backup routes.

1. The maximum channels the network offers = (16 x 32) + (8 x 8) = 576.
2. Total links to carry a single demand = (mean primary hops + mean backup hops)

   Total links to carry a single demand = (2.0 + 2.5) = 4.5.
3. Total demands carried = total network channels / mean hops per demand

   Total demands carried = 576 / 4.5 = 128.
4. Hence the Revenue Potential of the network is 128 channel demand units.
5. If the network provider charges customers £1,000 per channel,

   the RP of the network = 128 x £1,000 = £128,000.
6. In conclusion, the network must cost less than £128,000 to be profitable.

With such simple calculations it is possible for the network designer to estimate the load carrying capacity of network designs and compare the merits of a number of similar designs. Though this method is based on approximations, over a large network the mean figures used are likely to be sufficiently representative to allow rapid analysis of the viability of each design. They also allow a rapid ranking of each topology on the basis of likely revenue against cost, i.e. commercial viability.

A problem with this calculation is that the number of hops between nodes of a proposed network topology might differ from those in previously monitored topologies. A solution is to use the mean number of hops per route for the traffic demands allocated in the new design under review. These figures should be available from the routing report of the design software.

This has the advantage of being pertinent to the topology under review but suffers from potential inaccurate or unrepresentative mean values due to the small sample size, i.e. a single topology. In order to mitigate the need to use hop figures relevant to the new design, against the risk of unrepresentative figures due to the small number of routes, it is possible to base the RP figure on a weighted sum of the two mean hop figures. The hop figures are taken from both the current, new design and a historical record of all previous designs with the weightings derived from the sample sizes.

$$RP = \frac{K.\sum_{i=2}^{N}\sum_{j=1}^{i-1} C_{ij}}{\frac{S_1}{S_1 + S_2}.H_1 + \frac{S_2}{S_1 + S_2}.H_2}$$

where K = channel cost charged to customer;

C_{ij} = capacity, in channels, of link from node i to node j;

H_1 = mean number of hops from design method 1;

S_1 = sample size of method 1, ie number of routes allocated;

H_2 = mean number of hops from design method 2;

S_2 = sample size of method 2, ie number of routes allocated.

4.2.4. Resilient Path Traffic Allocation

In order that capacity be allocated for use in the event of any single link failure it is not necessary to double the entire allocated capacity. Following each possible trunk loss the surrounding trunks will suffer a combination of traffic loss and gain [Agarwal 1989]. In order to evaluate the total capacity required to carry traffic following any single link failure it is essential that the loss and gain of routes on each link is evaluated for all link failures.

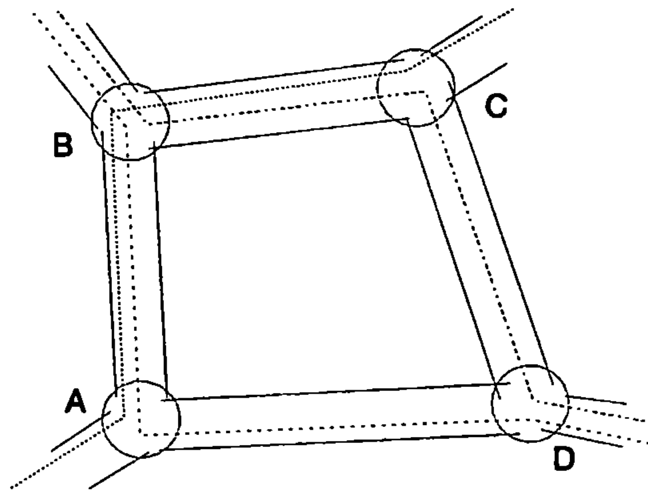


fig 4.8 Showing how link failure causes route displacement

From the diagram it can be seen that if link BC should fail then the DCB route will be lost and will need to use its backup path DAB. Consider link AB when link BC fails, it will lose traffic due to the ABC route being lost but will gain traffic due to the DCB route swapping to its backup allocation DAB. This failure analysis must be performed on all links for all other possible link failures.

Any single link must be able to provide the capacity to handle the maximum increase in traffic flow due to the failure of the worst case link. The maximum spare capacity that must be allocated on any one link will occur when the least traffic is lost due to a failure and the most backup traffic is routed over it due to the same failure.

If the maximum flow to be rerouted through link j by the failure of link i is defined as $g^{MAX}ij$ and the minimum flow to be lost on link j by the failure of link l is defined as $g^{MIN}jl$ then if G_j is the maximum provision of capacity necessary on link j,

$$G_j = \max [\, g^{MAX}ij - g^{MIN}jl \,]$$

The maximum capacity requirement under failure conditions depends upon the topology of the network and the routing of all connections using the failed link. This is potentially a complex calculation for networks of more than a dozen nodes and can only be undertaken when the entire routing of all primary and backup loads is known. However, it is necessary to allocate capacity for backup provision during the design phase of the network when links are being added and capacity determined. During the design phase it is not possible to make the G_j calculation since backup paths may not be established until the topology is complete. Any spur in the partial design offers no secondary paths if any component trunk should fail. In chapter 6 a method is developed to allow the approximation of requirements in the interim design phase before final capacity allocation and tuning is made.

The capacity allocated for resilience on a given link is not required for the failure of that link but those around it. It is therefore possible to have links with very low primary traffic capacity allocation but large resilient capacity allocation. This is most likely to occur on links connected to those heavily used for primary paths. As a network topology becomes heavily meshed, there are more alternative paths available to carry disrupted traffic in the event of a single link failure. If a greater number of links are available then the burden of carrying rerouted traffic can be shared among all those available as a result of the many more alternate paths that may be established. This has a correspondingly lower total backup capacity requirement for the network.

4.3. Components of a Network Model

In order that any form of machine based analysis of network topology can take place a model is required to describe the major components of the physical network and the traffic demand. The basic model may be manipulated and modified by chosen numerical methods to create a representation of the selected trunks and the routes allocated to all traffic demands. From this representation of the physical network it is then possible to calculate metrics describing the merit of the design based on cost, path lengths and connectivity for example. The network model forms the basis of all network design systems.

4.3.1. Traffic Demand

Traffic demand may be specified in terms of packets per second or erlangs for a packet switched network. The distinction between the two being made according to the network resource being utilized. Erlangs are used where network access ports are limited, the traffic arrival rate and call length determine the port utilization and the probability of gaining access to a port. Packet length and arrival rates are used where trunk capacity or node processing power is limited. Where network models are required to estimate mean queuing delay for packet switched networks, usually using the Pollaczek-Khinchin method [Frank, Frisch and Chou, 1970], [Kleinrock, 1970], [Tobagi, Gerla, Peebles and Manning, 1978] it is necessary to supply the probability distribution function for both packet length and packet arrival rate.

However, permanently assigned circuits are provided by AT&T ISTEL's TDM network. This research work bases all demands in terms of multiples of permanently allocated 64kbps channels. The NET TDM network can allocate channels of capacity 300bps to 1536kbps though AT&T ISTEL predominantly sells channels of 9.6kbps and 64kbps with the 64kbps being the most common. The restriction of demands to multiples of 64kbps channels is purely for speed, since a PC based analysis system will work fastest using integer numbers in the range 0 to 255, (one byte of memory) the basic unit of demand is the channel. Were more processing power available demand would be recorded in bits per second, bps, and give more flexibility to the model.

The node to node traffic demand, in channels, is represented by the 'demand matrix', a two dimensional array. The demand matrix may hold differing traffic requirements for node A to B and B to A, alternatively the requirement may be symmetric where A-B and B-A values are identical. This is normally the accepted case since channels in circuit switched networks are configurable for symmetric bandwidth between the originate and answer ports. Symmetric demand is assumed throughout the rest of this thesis.

The traffic demand that is the subject of this design work requires that the demand be satisfied under single link failure conditions and therefore primary and backup routes must be allocated. The demand between the i^{th} and j^{th} node is written:

$$\textit{Initial Demand} = D^{I}_{ij}$$

A system may be implemented for differentiating between the demands on the primary and backup routes in order that they may be allocated separately. This is done by initialising two demand matrices, one for the primary demand and one for the backup demand.

Where a demand matrix entry is zero no route will be allocated for the source destination node pair referenced. In its implementation the design/optimisation system will start with an initial demand matrix and attempt to find routes for all those demands with non-zero entries. When a demand has had a route allocated that demand is deemed to be satisfied and may therefore be zeroed in the demand matrix. This is done for both the primary and the backup demand matrices.

$$\textit{Primary Demand} = D^{P}_{ij}$$
$$\textit{Backup Demand} = D^{B}_{ij}$$

4.3.2. Network Link Cost Function

The network link cost function is one of the critical factors in the determination of network topology. Within a single country it is common for a general formula to be applied by the local PTT company to calculate the cost of a link of a given distance for a particular capacity. The cost will typically be broken down into a fixed setup charge to cover the installation of the link plus a further cost for each kilometre of the link. The formula may also have a capacity dependent term, in order that the cost of differing capacities may be calculated.

So that the cost of all links may be modelled a three dimensional matrix is required to hold the cost of the link between each node pair for each available capacity. The expression shown below requires a term to indicate the presence of the link from the i^{th} node to the j^{th} node in order that the cost be given as positive and real for allocated capacity or zero if no link is present.

In constructing the general expression for the link cost it is necessary to cater for the three cost components. Assume Ω_{ij} represents the sum of primary and backup channels to be allocated between i and j;

i. the setup cost that is related to the one off circuit installation cost;

ii. the distance related part of the cost, PTT charges include a cost per kilometre; let x_{ij} be the distance from node i to node j;

iii. the scale factor that makes the cost proportional to the number of channels provided on the link;

iv. let $f(\Omega_{ij})$ be the capacity dependent cost multiplier.

If c_{ij} be the cost per channel of a trunk between nodes i and j;

$$c_{ij} = l_{ij} \left[k_0(\Omega_{ij}) + x_{ij}\, \Omega_{ij}\, f(\Omega_{ij}) \right]$$

given

$$\Omega_{ij} = (\alpha_{ij} + \beta_{ij})$$

where l_{ij} is a [0,1] matrix representing the presence of a link between i and j;
if $\Omega = 0$ then $l_{ij} = 0$;
if $\Omega > 0$ then $l_{ij} = 1$.

All cost functions for real channel allocations exhibit a tendency for the per channel cost to reduce as the number of channels on the link increases - a form of bulk discount - hence the $f(\Omega_{ij})$ scale factor always has a negative, or zero, slope. This form of capacity, cost relationship is due to the reduced costs to the PTT of supplying the larger link capacities. The cost savings are the result of the PTT requiring less capacity division equipment between its high capacity national core network and the end user site to reduce the very high core bandwidth to the required user channel bandwidth.

A typical continuous function might be of the form:

$$f(\Omega_{ij}) = k_1 e^{-k_2 . \Omega_{ij}}$$

where k_1 is a scale factor on the cost function magnitude;

k_2 scales the rate of change of the capacity cost reduction.

This shows that the cost falls rapidly and then tails off for increasing capacity.

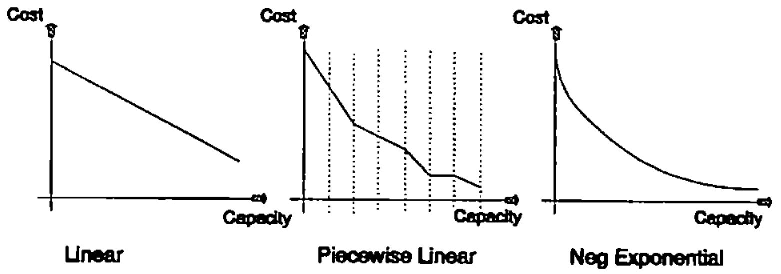


fig 4.9 Typical cost verses capacity functions

Balakrishnan & Graves [1989] use a concave cost function that reflects the realistic cost functions of network capacity providers by using a banded cost scheme, the per channel costs reduce as the number of channels required increase.

The cost function therefore uses a range value r which is determined by the number of channels required, w,

$r = 1$ for $0 < w < W_1$;

$r = 2$ for $W_1 < w < W_2$;

$r = 3$ for $W_2 < w < W_3$ etc.

$$C_{ij} = k_1(r) + k_2(r) \cdot d_{ij}$$

where k_1 and k_2 again represent the fixed and per unit distance values (k_1 may be constant for all r).

Taking the example given by Balakrishnan & Graves, the objective is to minimise the network cost, based on a concave function using the cost banding scheme.

$$z^* = \min \sum_{(i,j)} \sum_{r=1}^{R} c_{ij}^r \left(\sum_{t=1}^{T} x_{ij}^{tr} \right) + \sum_{(i,j)} \sum_{r=1}^{R} k_{ij}^r \, y_{ij}^r$$

where k^r_{ij} is the setup cost of link ij for capacity in the r^{th} band;

c^r_{ij} is the per distance cost of link ij for capacity band r;

x_{ij} is the traffic flow on link ij;

y^r_{ij} is a 0,1 variable representing whether the capacity is within the r^{th} band.

4.3.3. Route Selection

The selection of routes to satisfy each traffic demand requires that each is stored for later appraisal and reporting. In a similar form as the traffic demand matrix, the route matrix requires a list of each node visited along each path. The route matrix is therefore a three-dimensional structure.

4.3.4. Trunk Capacity Allocation

When routes are selected the traffic must be allocated to each link in the route and therefore a two dimensional matrix is needed to hold the number of channels allocated to traffic on each trunk between each pair of nodes. When costing a network topology the number of channels allocated is referenced to the link cost matrix in order to determine the cost of each trunk.

If separate trunk capacity matrices are maintained for primary and backup path allocations it is then possible to quantify the costs of primary and backup paths which is often useful in determining whether the backup is cost effective. There may be occasions where certain backup paths appear inordinately expensive in comparison to the majority of others. Where great disparity is found it may be possible to redesign the topology to minimise individual route costs. This type of decision is typical of design perturbation phases [Frank and Chou, 1972], [Gerla and Kleinrock, 1977], [Pierre and Hoang 1990].

4.4. Network Design Problem Formulation

Once a network model is established all problems are formulated along broadly similar lines, an objective function is written down that defines the aim of the optimisation. In most cases the minimisation of cost of the network is the objective which is then subjected to a range of constraints. The precise nature of the constraints applied are dependent upon the type of network under construction. Time delay constraints are only applicable to queuing networks [Whitt 1983] such as those of delay limitation. Some constraints are highly specialised such as those used in the prioritised link pre-emption scheme by the AT&T EINOS system which require the specification of links that may or may not be rerouted after disconnection following trunk failure [Agarwal 1989].

A typical cost objective function is usually written to include the costs of the network links, in the form:

$$z = min \, [\sum_{i=2}^{N} \sum_{j=1}^{i-1} c_{ij} \, l_{ij}] \;\; ;$$

where c_{ij} = cost of link ij;

l_{ij} = 0,1 indicates link ij is present.

The overall cost being the sum of all costs of each link from node i to node j, for all i and j, (j not equal to i). Constraints applied to network design methods may be split into four groups:

i in packet switched networks the first is the time delay constraint, (typically 200mS for the ARPA network and those based upon it [Frank, Frisch & Chou, 1970]);

ii the second being those specific to the problem that define such things as channel priorities and failure operation [Agarwal, 1989];

iii the third to control the network flow, i.e. to limit flow to values within the link capacity available, or to ensure capacity allocated is always utilised at a level greater than a lower bound;

iv the fourth to control the values feasible solutions take on, i.e. non-negative, integer and below maximum bounds.

4.5. Network Design Methods

Sharma [1988] provides an introduction to the general network design problem in a paper intended to illustrate the difficulty found in designing a usable computer network based on megabit capacity links. The paper is aimed at medium sized network managers in the corporate environment and assumes a sparse mesh network is to be designed. The author considers a limit on number of hops to be allowed in routes and specifies that at least two alternate paths are available for backup. A least hop route search algorithm is advocated (but none specified). The addition of 24 channel 1.554Mb/s (USA) T1 links is suggested until all capacities are sufficient and a redundancy (no description) method cannot remove any. The writer admits this is 'tedious' and suggests the reader cycle through the problem by hand until some form of minimum cost is found. No reference to any optimality is made but the reader is encouraged to repeat the perturbation phase seeking a cost minima. This serves well to outline the general approach to network design which can be greatly enhanced by the use of computer methods.

Following the formulation of a problem the solution must be attempted in an efficient manner. The outline for all general network design methods detailed up to this point is:

i. generate an initial solution;

ii. perturbate the design, use some form of link analysis, delete links judged to be poor or expensive, add links judged to be cheap or 'good';

iii. repeat ii. until the resultant objective function is within the specified limit.

The details of individual methods are interesting because the emphasis is usually on the generation of many random, feasible initial solutions and the creation of a range of viable perturbation methods [Kleinrock 1970], [Frank and Chou, 1972], [Gerla and Kleinrock, 1977], [Pierre and Hoang 1990].

Gerla and Kleinrock [1977] propose the following method, called the *Minimum Link Algorithm*:

i. determine all minimum shortest hop feasible paths for chosen j-k route;

ii. choose lowest utilised path;

iii. route required γ_{jk} along selected path;

iv. repeat for all j,k.

This is an efficient and fast algorithm, but unfortunately it is described as being insensitive to link loading and queuing delays and therefore not optimal under heavy traffic conditions.

Pierre & Hoang [1990] aim to use heuristics to 'drastically reduce the search space of candidate topologies'. They consider two level hierarchical network, with distinct access and backbone sections, very much in the same vein as the work carried out in the rest of this thesis.

The method involves first generating an initial solution:

i. given a k-connectivity requirement select links (no method given - no discussion of optimal search, no discussion of link selection criteria);

ii. perform routing (shortest path);

iii. perform link flow assignment;

iv. perform link capacity assignment;

v. compute link costs, mean time delay.

The authors suggest the initial solution has a worst case complexity $O(N^4)$ which is not necessarily relevant given the manner in which the process was conducted (i.e. ordered routing, flow, capacity assignment, no iteration between phases) and the fact that the quality of the solution is unknown.

The basis of the system is a rule base of perturbations that consists of types such as: perturbation selecting rules; capacity modification rules; link addition rules; link deletion rules; link substitution rules and positive or negative example defining rules. The rules are fairly simple and certainly intuitively obvious in their application. This system makes use of an expert system based on a forward chaining inference engine. Examples are given for 2,3 and 4 connected networks.

Saksena [1989] notes the importance of hierarchical network structures which are likely to become increasingly significant as networks grow substantially over the coming years. He states that the number of hierarchical packet-switched network analysis papers presented to date is very low (two or three) and that as networks get larger this is the next logical step. He attempts to produce a range of design guidelines for such large scale hierarchical networks.

Hierarchical routing requires that higher level nodes are not allowed to switch paths down to lower nodes and then back up again. The routing is determined by hierarchical clusters, which define the nodes to which routes connect. Saksena also notes that greater hops for routes results in greater network resource usage, hence likelihood of greater overall capacity requirement (cost). The author looks at the methods for reducing network capacity, by the elimination of low utilisation links. He gives a good demonstration of how combining under utilised links can reduce the overall cost for same delay figure. This work offers useful assistance to the appreciation of perturbation design methods as shown later in this thesis.

The methods given are shown to generate 'near-optimal' networks, based on an initial solution which has all direct routes selected. The link utilisation is then examined and decisions are made as to whether to eliminate the link and re-route paths travelling through it. A number of iterations are performed to reroute and assign the required capacity to meet recalculated delay constraints. The networks tackled by the author had between 10-40 nodes which represents a fairly substantial number for a large single vendor network - not dissimilar in size to that of the AT&T ISTEL network. A range of traffic demand matrices were tried with symmetric and asymmetric loads. This paper is significant in that it is one of the very few (possibly only) one to **set hop limits for routes**. This is contrast to the majority of work that uses the k-connectivity limit. An important feature of the method employed is a **figure of merit** which is defined as a ratio of the route cost relative to the direct link cost.

Singhal [1989] outlines a number of methods, among which **simulated annealing** is the most interesting. It randomly samples points in the design space (by performing random perturbations) and tracks the best solutions, the space over which jumps are made (i.e. the size of the perturbations) is reduced. This is analogous to slowly reducing the temperature in a metallic medium to allow crystalline growth to occur - hence the term 'annealing'. The system is acknowledged as only finding near optimal solutions but has been shown to work well, producing good results within short timescales. An important feature of this method is that it is able to (allowed to) move the cost function uphill to move out of local minima, it is therefore said to be better than greedy algorithms that look for the fastest improvements in the cost function. This property is considered important and is applied in work described in the following chapter.

Yokoyama, Miyake & Nakajima [1988] use a backbone design that appears to be based on a multi-level hierarchical clustering system, where the hierarchy is determined by node location, cost and traffic volume. Link capacity is minimised until delays reach predetermined limits. Lagrangian methods are used to iterate network cost. No details of the problems formulation are given but the method outline gives some clues to the factors the authors consider important in network design.

4.5.1. The Capacity and Flow Based Problems

A distinct class of network design and analysis methods are based on the allocation of capacity to maximise flow between nodes. The objective of such design systems is to minimise the allocation of capacity in order to minimise network cost while maximising the flow of traffic on each link in order that the minimum unused capacity is incurred.

The capacity flow assignment, CFA, problem [Gerla & Kleinrock 1977], [Gavish & Neuman 1989] is associated mainly with packet switched networks. A range of traffic demands are known or estimated for a set of nodes and the capacity of links between them must be assigned such that the expected queuing delay of the network is within a predefined limit. The objective of such systems is to minimise total link costs, the constraints of the problem are devised to ensure that queuing delays are within bounds and all flows are valid.

Capacity Assignment requires the statement of the main objective function which usually describes the total cost of the network which is made up of the local user to concentrator link costs, concentrator to centre cost plus the total concentrator cost.

4.5.2. Linear Programming

There are many mathematical problems that require the determination of the optimal values for a range of parameters subject to a number of constraints. Linear programming (LP) is a method for solving problems, with linear variables, based upon the formulation of a mathematical model which accurately (or to a level sufficient for the purposes of the particular application) describes the relationships between the constituent variables of the problem. The aim of linear programming is to optimise a specific parameter (or sometimes more than one) of the model which is a function of a number of other variables. Typically, in many commercial applications, a minimum cost solution is desired and hence the objective function is a measure of cost.

The fundamental equation written in forming a linear program is the objective function, for which a minimum or maximum value is sought (if it exists or can be found). The objective function is subject to constraints which limit the acceptable values of the individual variables. Constraints are written in the form of inequalities to bound the range of any variables where appropriate. Linear programming is a subset of mathematical programming where the relationships between all variables are linear. The most widely used method for solving LPs is the Simplex method developed in the late 1940s by George Dantzig [Moret & Shapiro 1991] and widely used in network design systems, [Balakrishnan & Graves 1989], [Mirzaian 1985], [Mulvey 1978], [Singhal et al 1989] though a new high speed (and complex) method has been produced by researchers at AT&T called the Karmarker algorithm [Cheng et al 1989].

In the network design field all numbers that are dealt with take on integer values - link capacities and traffic demands may be measured in terms of 'channels'. The implication of real solutions being limited to taking on integer values is that any continuous variables would require adjusting to the nearest integer. Integer programming (IP) is a subset of linear programming and the satisfaction of the requirement to produce integer results leads to rounding and, as a consequence, potential difficulties in finding unique solutions. It is possible for integer programs to have solutions that may not be found within a polynomial time function of the primary variables - for large problems no solution can be found in useful execution times without resorting to using heuristic methods to approximate values.

The vast majority of network design systems [Kleinrock, 1970], [Frank and Chou, 1972], [Boyce, Fahri and Weischal, 1973], [Gerla and Kleinrock, 1977], [Mulvey, 1978], [Agarwal, 1989], [Gerla, Monteiro and Pazos, 1989], rely on integer programming methods for creating feasible initial network designs, the inherent difficulty in finding integer solutions requires that perturbation schemes are then used to improve upon the design and further minimise the cost.

The perturbation schemes employed are not specific to integer programming, but are a general class of methods for examining design metrics and making decisions on which trunks may be added and deleted. The perturbation schemes are what may be described as threaded decision processes and are discussed further in the following section 4.6.

4.5.3. Formulating an Integer Program for Network Design

During the early work on this research project attempts were made to formulate an integer program to facilitate the design of a computer network conforming to the particular constraints of the design required, for the AT&T core network these are:

i. limit on number of hops per route;

ii. provision of link and node disjoint primary and backup routes.

The formulation was started with the simple objective of minimising the cost function. Due to the banded costs for UK and European link capacities the formulation was similar to that of Balakrishnan & Graves [Balakrishnan & Graves 1989].

The objective is to minimise cost (based on a concave function):

$$z^{*} = \min \sum_{(i,j)} \sum_{r=1}^{R} [c_{ij}^{r} F_{ij} + s_{ij}^{r}] \; y_{ij}^{r} \;\; ;$$

where s^r_{ij} is the setup cost of trunk i to j for capacity in the r^{th} band;

c^r_{ij} is the unit channel trunk cost between nodes i and j for capacity in band r;

F_{ij} is the flow on trunk ij due to all routes;

y^r_{ij} is a 0,1 variable indicating whether link capacity is in the r^{th} band.

A series of constraints were developed to control the link flows and the routing hop limits. A route vector is defined between the nodes x and y as consisting of a maximum of N nodes in the path, the first node is x and the N^{th} node is y.

$$R_{xy} = \{ x, r_1, r_2, \dots r_N \}$$

The flow on each link ij is F_{ij} and can be found by summing the total traffic due to routes from x to y, for all x and y, using the link ij:

$$F_{ij} = \sum_{x=2}^{N} \sum_{y=1}^{x} (D_{xy} \cdot J(x,y).R_{xy}) \quad ;$$

where D_{xy} is the demand for the route from x to y using trunk ij;

R_{xy} is the route from x to y;

J(x,y) is a 0,1 function that indicates whether the route xy uses trunk ij.

However it became evident that the problem was becoming increasing complex and the processing power available to this project had to be taken into account. Much of the network design work in this field is performed using IBM mainframes [Bahl & Tang 1972], [Boyce, Fahri & Weischal 1973], [Danzig et al 1979], [Gerla & Kleinrock 19¢77] and newer specialist high performance UNIX workstations [Cheng et al 1989]. Under budgetary constraints the processing power available was due to an IBM PS/2-70 personal computer. Using commercially available Linear Programming packages it would clearly not be possible to solve the model under development.

With the limitation on processing power and the evident size of the network design problem it was therefore necessary to seek to develop a system that is suitable for application on a personal computer and heuristic methods must be sought in order that good solutions may be found in humanly useful computer execution times.

4.5. Network Design Systems

There are a number of network design systems commercially available for the design of TDM core networks. Norman [1992] briefly describes a series of commercial products, all supplied by American companies and consultancies. BESTnet from BGS Systems Inc [Waltham Mas, USA], COMNET II.5 and NETWORK II.5 from CACI Products Co [La Jolla, Ca, USA], NetConnect from Connections Telecommunications Inc [W. Bridgewater, Mas, USA] and NetMaker from Make Systems Inc [Mountainview, Ca, USA]. It is clear that the majority are simulation or analytic systems to estimate capacity utilization and delay performance for packet networks. With the exception of NetMaker it would seem that analysis of user supplied topologies is intended to support the network designers own judgement in modifying the topology to improve performance or support new traffic, they do not offer 'green field' design.

The primary 'design element' of these packages is a trunk tariff database from which the network designer is intended to cost and deduce network solutions shown to be viable from simulation or analysis. An exception to this is the Mesh Linker system from Quintessential Solutions Inc. [San Diego, Ca, USA] which 'based on predefined traffic and controls provides optimized backbone networks that interconnects all the concentrators'. Unfortunately the only commentary on the design methods employed is 'this design tool is based on proprietary heuristic analytic techniques'.

Given the high commercial value of good design systems the lack of technical detail is not surprising, though a design system for which details of a true green field ('dessert', US parlance) design method have been revealed (in brief) are those of NetMaker from Make Systems Inc.

4.6.1. Make Systems Inc, NetMaker

NetMaker is an expensive package, costing approximately $100,000, developed and marketed by Make Systems Inc. One of its primary features is a network design system for circuit switched networks, including the device specific methods for the NET IDNX hardware installed by AT&T ISTEL throughout Europe. The precise design methods are commercially sensitive [private correspondence with Make Systems Inc] but in principle work as follows:

i. a minimum spanning tree is constructed between all nodes;

ii. all end to end traffic requirements are routed and capacity selection of trunks performed;

iii. to meet single link failure resilience criteria, the effect of losing each link in turn is modelled and the cheapest alternate path sought;

iv. perturbate by, testing for reduction in network cost by:

   a. delete trunks with successively lowest utilisation;
      allocate cheapest new trunk/s to carry lost load if necessary;

   b. delete successively most expensive trunks;
      allocate cheapest new trunk/s to carry any lost load.

NetMaker is not claimed to produce optimal network designs but is intended to produce costs within a target of 3% of optimal. It must be noted that no statement of expected optimality or any expression of confidence in the design is available from the package.

There is no real explanation for the use of the MST as the starting point, indeed, with reference to the limitations of such threaded design systems, it would appear that this poses a severe limit on the possible outcomes of the design process. Since NetMaker uses a number of perturbation schemes it is essential that they overcome any excess capacity problems of the MST starting point and any resultant path 'lock-out' problems [see section 2.7.4]. The MST is chosen for its minimal link distance property yet this is only a valid objective for a connectionless network topology, where no dedicated paths and capacity are allocated. A problem will immediately occur if any high traffic demand from one end of the topology to the other should exist since they would, in principle, merit a direct path. It is the lack of sensitivity to the needs of a particular traffic profile between nodes that leads to some concern over the starting point of this method.

CHAPTER 5

New Network Design Processes

5.0. Introduction

This chapter describes the manner in which the new computer network design and optimisation method was constructed. There are many ways of tackling this problem and the critical factors are discussed. The problem's complexity was investigated to determine the features that lead to the most significant computational problems and identify areas suitable for speeding up by the use of heuristic techniques. The use of analysis methods normally applied by human network designers was examined and a set of rules and understandings of network design requirements were developed. The basic building blocks required by a computer based design method suitable for applying the new methods are discussed. The development lead to a framework for a complete functioning network design system. An emphasis in this work was placed upon the creation of a practical design tool suitable for use in real design situations. Much of the work is specific to circuit switched core network design for AT&T ISTEL; however the methods would apply equally well to a number of similar situations where trunk selection and capacity allocation are required for a known end to end traffic demand.

In considering the design of large scale international communication networks based on circuit switched cores the largest capital investment is made in providing high capacity national and international data communication links. The design process must therefore concentrate on reducing line rental costs while meeting the node to node traffic requirements of all customers. From a financial point of view the cost of equipment is counted as a capital asset and may be devalued at 20% over five years, the 20% of purchase price therefore appears as the year on year cost of the equipment. However, international trunk rental costs are charged at a fixed monthly rate, usually on five year contracts. To compare, a £100,000 node appears as a £20,000 annual cost, depreciating to zero after 5 years, whereas a line rental of £30,000 per month costs £360,000 per year and can be seen as a £1.8M commitment over 5 years.

There is little scope for reducing node equipment costs since the locations of nodes are fixed as part of a business plan for the network. A particularly large localised growth in traffic would be required to justify a new node site. In an international network there is a compelling need to offer nodes located in each major capital city, primarily to capture the majority of business associated with corporate headquarters usually found there. The sale of network services to new customers can then be made using short local connections (tail circuits) to the core network which provides the international communication. However this is just one small part of the larger commercial issue behind network design that must be considered by international suppliers. The work presented here will concentrate on the primary objective of reducing line rental costs and satisfying the traffic demand criteria.

5.1. Network Design Complexity

The true definition of the 'topology' of a network is a description of its geographic locality [Oxford Dictionary]. This translates into a detailed description of the sites and connections between nodes, including their characteristic parameters such as number of inputs, outputs and link capacity. In order to meet a reliability criterion a minimum number of paths to each node might be specified. The specification of the number of paths between nodes may be made in one of two ways: either wholly or partially distinct paths must be made available. Two paths that are entirely distinct, i.e. share no common links, are referred to as 'link disjoint'. This is a more strict criterion than merely requiring a number of alternate paths, whether or not they share common links. For large scale network design there are two basic design stances that may be taken. The first would be to assume that should a link fail the 'outage' will be short and any lost traffic is dispensable, there would then be no need to ensure alternate paths between nodes. The second stance is to assume failures are likely. If the second stance is taken it is necessary to determine the number of concurrent failures that are to be allowed. For increased network availability it is necessary to provide an additional capacity contingency to cater for a greater number of simultaneous failures and support more alternate paths. This leads to a greater resultant network cost and cost to the customer.

The variability of topology arises due to a number of factors, but principally because there are a great many ways of connecting N nodes together. In order to appreciate the computational complexity of the link selection procedure in the network design problem an estimate of the number of distinct network designs is made. The design variables are based on the total number of combinations of node to node trunk connections and the range of possible link capacities that couple with all the possible routing schemes for all possible loads between each node.

The number of possible mesh networks can be found by calculating the number of permutations of links in a single link failure resilient configuration between N nodes. The minimum number of links occurs in a ring and maximum number in a fully connected mesh:

for a ring the minimum number of links $L_{min} = N$;

for a fully connected mesh the max links $L_{max} = N (N-1) / 2$.

For a network design problem the number of topologies that require testing for feasibility is based upon the number of permutations of L links between N nodes for all values of L between the maximum and the minimum. A sub-mesh is considered as any mesh network with L links, less than or equal to the number in a full-mesh:

the number of sub-meshes $M = L_{max} - L_{min} = (N^2 - 3N)/ 2$.

If a purely link based topology design system were employed, to determine which links to select, it is necessary to consider all combinations of L links within the M possible submeshes. However, it is also necessary to test whether the topology is feasible and can support the required load. This requires the operation of a routing algorithm which itself has a complexity related to the number of nodes and connected links. Since the number of combinations of trunks in each topology is related to the factorial of the M term this number has a rapid rise with increasing node count.

Ideally a scheme would be used where the capacity is automatically selected to accommodate the precise number of channels required after the link and route selection. This method would be valid if channel capacities exactly fitting the requirement were available, but unfortunately only a small range of capacities are commercially available. The selection of specific capacities, as required, is therefore not possible and compromises must be made. This leads to a need to iterate through the design process, searching for a best fit of requirements against available trunk capacities. This potentially lengthens, without bound, the search process in a similar manner to that of integer programming when compared with linear programming.

In addition to the trunk selection problem, the routing complexity is related to the number of combinations of available links for constructing a path. In order to emulate the NET TDM routing parameters, a limit is placed on the number of links that may be used in constructing a route between nodes. The maximum number of hops is written H_{max}. This limits failure probability, total network resource requirements and limits time delays. Even in TDM networks, where delays are minimal (4.1 mS per node for NET IDNX Technical Specification), delays can be significant. For multiple hop paths the fixed delay in addition to the propagation delay can cause problems to customer circuits. A network provider will usually operate against a written statement of maximum transit delay which limits the maximum number of hops.

The number of possible routes is therefore related to the number of combinations of H_{max} trunks from the total trunks possible in the topology. For example the number of 7 hop paths between 20 nodes is 390×10^6 (7 permutations from 20) and the number between 21 nodes is 586×10^6 (7 permutations from 21).

The solution of the Network Design Problem, NDP, therefore requires the evaluation of a great many possible solutions and is inherently difficult. The 'hardness' of this problem has been proven in the strictest sense by Lenstra et al [1978] who demonstrate that the NDP is in a special class of problems known as NP-hard. In the mathematical sense an easy problem of basis N (N nodes in the NDP) may be solved in a number of operations proportional to a fixed power of N, eg N^7, N^{12}. A hard problem, by contrast, may take an exponential power of N to solve, e.g. 2^N. An NP-hard problem is one that may be solved in polynomial time using a non-deterministic algorithm, i.e. one that supports the feature 'goto A or B' where the choice is random but optimal. There are no known methods at present for creating such algorithms [New Scientist, issue 1851, 1992].

Since the complexity of NP-hard problems expands so rapidly in N the number of problems that can be solved in human time (i.e. useful in the life of a human) is very small, only suitable for low, though problem dependent, values of N, .

The routing of traffic from all nodes to all destinations cannot be achieved without knowing which links and capacities are available, however the links and capacities may not be chosen until the routes are known. It is therefore necessary to find a method for 'breaking the loop' and making 'good' estimates of optimal link positions and routes.

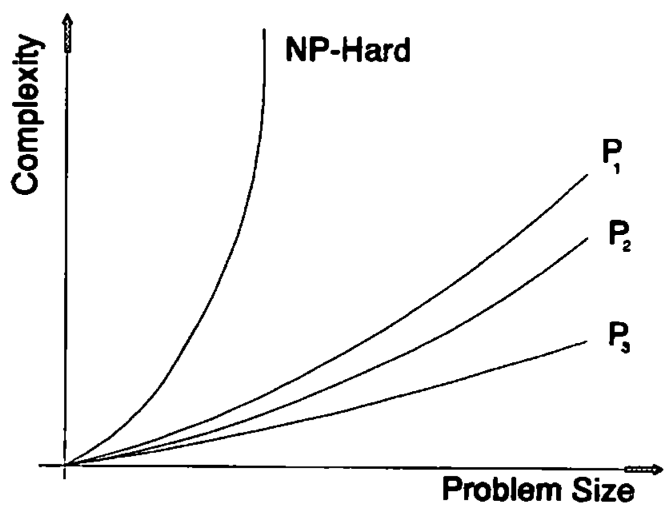


fig 5.1 Complexity for NP hard and P class problems

If the allocated links were known the routing problem would become easier. It is not clear that a point in the evaluation cycle exists when the links can be definitively selected. Attempts to solve large problems, such as those with more than a dozen nodes, cannot be made explicitly, within the processing and analysis constraints of current personal computer technology. In order to solve problems of a practical size it is therefore necessary to develop heuristic methods. The areas of the problem likely to benefit from the use of heuristic methods are in reducing the time to search all nodes for 'good' (low cost) routes and in controlling the allocation of capacity to create 'sensible' topologies, i.e. a number of the possible topologies would be evidently incorrect to the user. For example topologies with high capacity to nodes with little or no traffic and large numbers of small capacity trunks where tariff advantages make fewer large capacity trunks cost effective.

5.2. Developing a Network Design Method

Following some of the problems found with the formulation and analysis of linear programming models, in chapter 4, an alternative system was sought. Linear programming models, even when they can be formulated for a particular problem, offer a rigid and single objective search method for evaluating the vast range of possibilities. They have only the ability to search for the solution that minimises the overall cost (where it is made the objective). They cannot select solutions of slightly higher cost that exhibit better characteristics, for example lower mean route length (improved reliability), reduced number of trunks (minimising maintenance), or even distribution of spare capacity. Unless those desirable characteristics can be defined in the objective and constraints of the 'program' they will be ignored. It is questionable whether it is possible to define all 'good' network topology characteristics explicitly and the computation overhead may render the problem intractable. Large scale networks in Europe cost many millions of pounds in annual rental of internodal circuits and a portion of the topology is likely to represent strategic planning as the result of a selected 'market position'. In this case a minimum cost solution is the primary objective but additional market requirements may need defining in the form of complex constraints. A linear program must run until its completion to produce a single result that is definitive for the given inputs. It must therefore spend a considerable portion of its execution time examining infeasible or poor options which are then ignored. Should the problem be large and execution times lengthy it is not possible to interrupt the process and gain a partial solution of any use or known quality. The network design system requirements identified by AT&T ISTEL were based on the development of a system that would tackle large problems rapidly, designing networks based on the use of a structured search for the best components of a solution. The best components are those determined by metrics specific to large scale network design. A large contributor to the high speed design requirement is the commercial time pressures for producing cost analysis and business scenario evaluation. Strategic decisions may influence the design parameters and even traffic and node requirements. In designing large scale international networks the level of strategic capacity that may be committed to is dictated by revenue forecasts (or targets) for the network. In preparing a market strategy a budgetary analysis and financial planning are required to dictate the level of expenditure that may be allocated to network infrastructure expenditure. Hence, a large part of the design requirement is 'what if' modelling based on a range of possible traffic forecasts from a number of sales scenarios.

It is necessary to examine the potential of a software program that can 'intelligently' search for near optimal configurations of a network for a given set (or number of sets) of design requirements. The use of intelligence can be achieved by analysing the metrics that a human analyst might apply and then translating them into a number of rules for use by the design software. The resulting method should benefit from the heuristic metrics which can eliminate the need for vast numbers of unnecessary or fruitless calculations and hence lead to good solutions within a time-frame acceptable to the network design engineer.

These heuristic methods are developed from the analysis of the critical factors that influence the cost of a network, its ability to carry a given traffic requirement and its ability to return a revenue from carrying that traffic. In examining the desirable properties of a network a number of features are sought that may be described in a mathematical sense. A number of techniques may be required to eliminate the need for performing unnecessary sequences of calculations on topologies that the network designer would immediately recognise as poor.

It is necessary that the resultant design methods provide a sufficient timescale reduction over the brute search methods to make them viable in a commercial environment. It would be pointless to tackle a problem normally requiring 10 machine years of processing with a method that only offered a factor of 10 speed improvement. The resultant year to perform a design is just as commercially unacceptable as one taking 10 years. However a system requiring only 30 minutes to complete the design of the same network, producing results of a possibly, slightly lower quality than the year long method is entirely acceptable in most practical situations. It is not assumed that quickly created, poor designs are favourable, rather, a number of rapidly created alternative designs may be more useful. A number of possible alternatives may be analyzed or perturbated further by the network designer with the result that a good final design is arrived at within a matter of hours. In contrast to the requirement of linear programs that they run until completion, it is possible that heuristic methods can be executed with sufficient speed to produce partial solutions or a range of designs that may be comparatively judged by the designer. When presented with a number of possible computer generated designs he may then see the relative merits of each and judge the spread of costs and the savings possible following further analysis. The ability to produce multiple designs, under slightly differing conditions, is important when heuristic methods are employed since without this, there would be no way of judging the quality and effects of the approximations used. It is therefore seen as important that, for a fixed set of network design requirements, a measure of variability in the design method is possible to produce a number of alternative solutions. Since heuristic methods inherently search a subset of all possible topologies and do not (as yet) guarantee the optimal solution in short execution times, a variable method is desirable. This is because adjusting the search method offers a chance to encompass a greater possible area of the solution space and a greater likelihood of coinciding with the optimal solution.

With the rapid growth of large networks and the ever changing traffic requirements the fast approximating design process is entirely acceptable in view of the considerable time otherwise required to find the optimum solution. An additional complication is that a single 'optimum' design may have unacceptable features, such as low connectivity to key nodes or high spare capacity in isolated regions. A further design phase would be required to work around any preferences and the benefits of a slow optimal search are then lost. The growth of network traffic must also be taken into account since a small increase in demand in part of the network could potentially lead to a significant change in the topology. It is equally possible that the topology might change little with large swings in traffic demand and the sensitivity of the topology to changing traffic loads offers an important clue to how robust the design would be in a commercial environment.

5.3. Basic Elements of a Network Design System

A method to allow the analysis and design of a network requires not only a set of rules for making design decisions, but also a model representing the network and a number of basic functions to interrogate and manipulate the model.

The basic network modelling components: demand matrix; cost matrix and trunk matrix, have been described in chapter 4 but the supporting functions required by a network design system are now described. In order to perform topology design for a core network, software components are required to allow the basic functions of network synthesis and analysis to be performed in a manner compatible with the design control software.

The lowest level functions required are those that allow the control software to determine the status of the network model such as:

i. finding the capacity of a link between nodes;

ii. determining if a path exists between nodes using a number of intervening links;

iii. determining if a traffic demand exists between two nodes;

iv. determining the cost of adding or upgrading a link in the network;

v. determining how much network resource would be consumed in satisfying a traffic demand.

There is a requirement that the control software may alter the state of the model for operations such as:

i. allocating capacity on a link or number of links to a given traffic demand;

ii. adding a new link to the model;

iii. altering a design parameter used by the design software routines.

In addition to the basic manipulation and interrogation methods there is a need for software routines to handle basic network specific functions such as:

i. find a route between two nodes for the minimum network resource cost;

ii. report on the cost of the current state of the network design.

The designer using the software must be given a full opportunity to interrogate the model for analysis, reporting and user appraisal purposes, typically this will require:

i. list of all links allocated;

ii. list of all traffic demands allocated and those not yet allocated;

iii. list all routes allocated between nodes;

iv. cost of all links allocated;

v. utilisation figures on all links, i.e. capacity used against total capacity;

vi. list of how many hops are used in each route between nodes.

By implementing a range of query routines to offer a number of ways for analysing the state of the model at any stage, it is possible to build up powerful design methods. The query routines may take on the form of questions similar to those a human might use in examining a network design. For example, a query might determine all links with less than 75%

utilisation or find all routes using less than 3 hops. The basic query routines may be used in larger more complex design methods.

5.4. The Outline Network Design Process

It is possible to approach the design specification of networks in many ways: the simplest is to concentrate only on the topology and make successive link connections until the network guarantees some predetermined level of connectivity. This is typical of the 'k-connectivity', design criteria for feasible networks used by some design methods [Pierre & Hoang 1990]. This is only appropriate where traffic levels are small compared to the trunk speed or they are unknown. A typical design might require a connectivity of 3, each node being connected to a minimum of three others.

An alternative method for defining network capacity requirements involves the step by step construction of the network by allocating capacity to each and every load carried across the network. Since the provision of capacity through the network for each demand requires the development of routing methods, the analysis is very much more complex. The design complexity is a result of the need to evaluate the relative costs of all possible routing options for carrying each traffic demand through the network. There is then the problem of choosing an order in which to select each path where each selection may be influenced by all the preceding route selections. Since the capacity chosen to satisfy earlier demands will affect the perceived cost of each subsequent selection, each decision must assume the consequences of those previous and accept that it will itself influence all subsequent decisions. This is termed a threaded design process.

5.5. The Threaded Design Process

In assuming that a model of node positions is known and that a traffic demand matrix is given, the basis of a simple design method might be a repeated test for what is considered to be next best link or route to add. As with any such 'threaded' process, and stated by Moret & Shapiro

[1991] there is no guarantee that following a series of apparently optimal moves (a 'greedy' algorithm) the end point is an optimal outcome. As an illustration of this in the networking sense, consider a partially connected network. In order to evaluate the relative costs of a series of routes the current link costs are used for costing the total routes. However, as the design progresses the links will undergo a change in both their absolute and relative costs. At some time in the future a series of links selected for a particular route might transpire to be more expensive than others that have subsequently become more cost effective. The process is inherently sensitive to the time at which a decision is made. Using an analogy from the game of chess, the taking of the opponent's queen's pawn might be a high risk, low yield move during the early stages of the game, however, as the game progresses the risk may lessen and the possible reward grows.

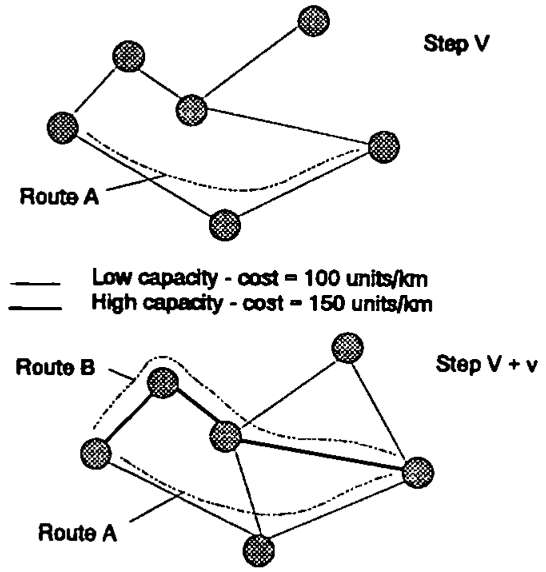


fig 5.2 The change in apparent route costs

The figure 5.2 above shows how the cost of route A appears cheapest at step V because the distance, and hence cost, of route A is lower than that of route B. However, at some future stage of the design process, V+v, some of the trunks have been increased in capacity and the cost per channel will have reduced. At stage V+v it is cheaper to select route B instead of route A.

This has a complex effect on the apparent cost of a link throughout an analysis and a simple case be examined by way of illustration. Consider the capacity allocated to a single link when required to carry two demands D_1 and D_2, the D_1 demand is allocated first at step n of the design process and demand D_2 is allocated at a later stage (n+m).

Consider a link from X to Y:

capacity 64kbps costs 100 units;

capacity 128kbps costs 150 units;

demand D_1 of capacity 40kbps;

demand D_2 of capacity 60kbps.

1. Allocate demand D_1:

   minimum link capacity to carry D_1 is 64kbps at a cost of 100 units per 64kbps;

   proportional charge to D_1 for capacity used = (40 * 100) / 64 = 62.5;

   cost per kbps = 1.56.

2. Allocate demand D_2 in addition to demand D_1 on link XY:

   total capacity = 100kbps;

   minimum capacity to carry load = 128kbps;

   proportional charge to D_2 for capacity used = (60 * 150) / 128 = 70.3;

   cost per kbps = 1.17.

It can be seen from this simple example that to allocate the second demand D_2 after demand D_1 has resulted in a lower cost per kbps of capacity. This effect is referred to as the time varying cost function.

If a network design procedure is carried out with static link costs (e.g. linear programming) then any decision to use a link incurs a fixed, known cost. However, if a threaded optimisation is performed then the costs of links changes as the traffic allocated is varied. A complication being that the incremental cost of channels may well not be linear; a predominant situation the telecommunications world. Hence the incremental cost of adding traffic to links varies as the design process progresses. It is necessary to examine closely the formulation of the cost function and understand the way in which it influences the selection of links at each stage of a design process.

5.6. The Threaded Design Decision Tree

The decision tree diagram can be used to illustrate the *inherent consequence problem* and implications of using threaded design methods.

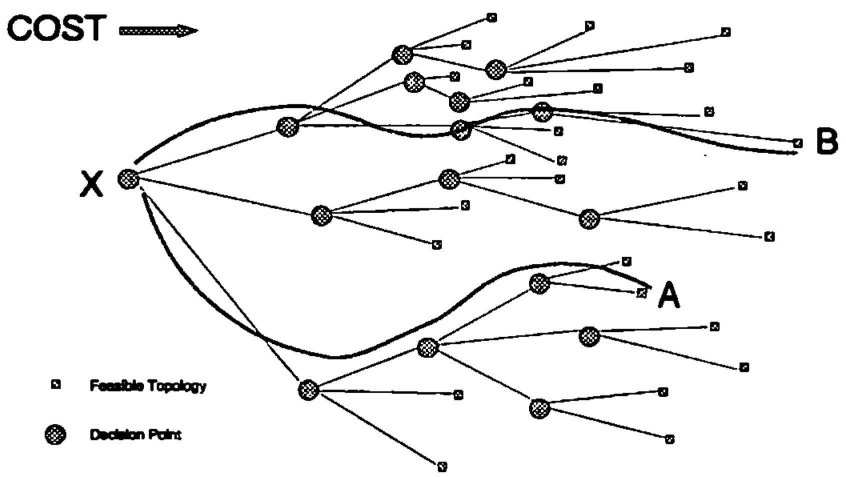


fig 5.3 Partitioning of the decision tree from early stages

For path A the decisions at point X leads to a best overall cost of 90. However, for the path B followed from point X leads to a best possible overall cost of 130 units. If point A represents the best possible topology then even the very first trunk selection on the path to topology B prevents the threaded design reaching the optimal solution directly. A method is required to allow the topology to 'jump' from one branch to another.

The tree structure only partially suffices in indicating the complexity of the threaded design process. In reality it is possible to reach many topological states from other pre-eminent states.

However this does not imply that the network is of an equivalent cost/capacity ratio. Since the paths are allocated in order to determine link requirements it is more than likely that different design attempts may reach the same partial topology but with differing traffic allocations to paths through the trunks. The best network then becomes the one with the minimum allocation of capacity for the greatest volume of traffic carried. This can be represented on a three dimensional tree where the third dimension is that of 'allocated capacity'. There may be partial topologies that are identical but have been reached by allocating different traffic demands, one may be more cost efficient that the other.

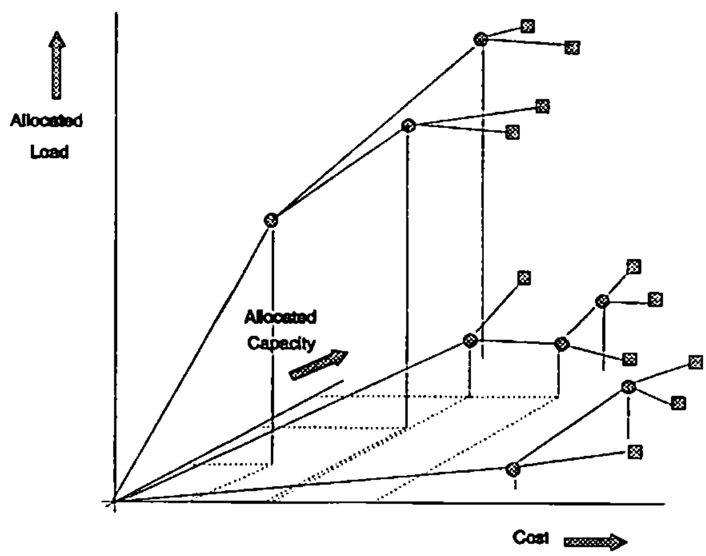


fig 5.4 Illustrating the three-dimensional threaded search through various allocation options

If a partial topology has been reached by two different series of decisions, though the trunk selections have coincided they are unlikely to have converged; subsequent capacity allocation decisions are not bound to be identical since the decision history is different. This point is illustrated below in figure 5.5.

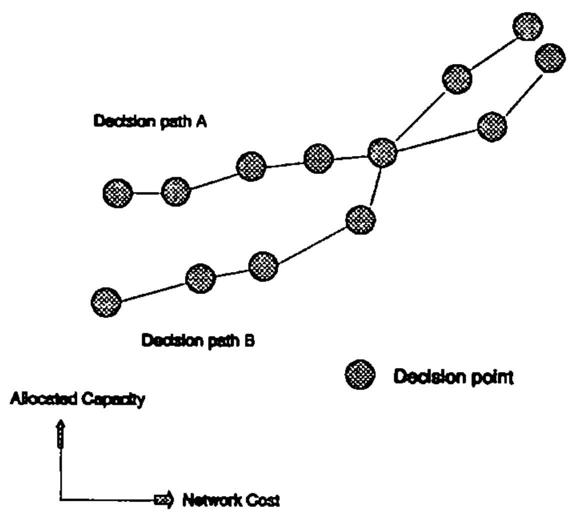


fig 5.5 Coincident intermediate topologies do not follow the same subsequent path

In order that the 'consequence' problems of threaded decision design methods may be addressed, the use of perturbation methods upon the final design ~~are~~ is required. Either a means is required to detect that the current topology is not on a branch that leads to the optimal solution or it may be accepted that the initial design method does not guarantee any optimal solution and the perturbation scheme is deemed essential. The first option would require the use of a lookahead method, where the effect of making a string of decisions is evaluated without committing to them. Since the decision making process is already likely to be computationally intensive, increasing the workload to contemplate the further permutations as a result of each decision is not attractive. Design times might lengthen to an impractical level. The perturbation scheme may be implemented as a clearly defined design phase, with a range of cost improvement heuristics. If they perform no cost reduction then the 'current' design is maintained. This has the benefit over lookahead methods of a potentially much lower computational effort, producing usable results at each pass. It does however place a demand for some ingenuity in the analysis of what constitutes a good perturbation scheme.

The effect of the perturbation operation upon the decision tree is illustrated in the figure below. A single link or number of links may be removed and the topology reverts to a lower cost state. This state need not be the one that has been a preceding topology in the design process.

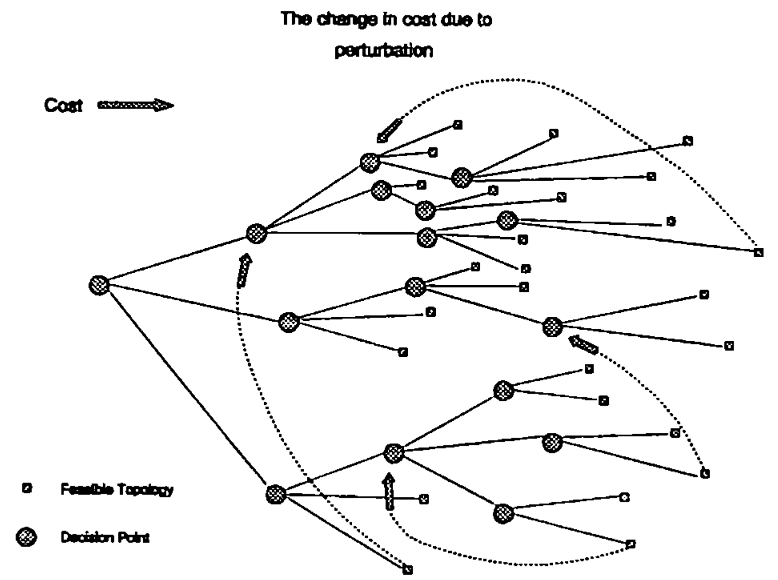


fig 5.6 Perturbation bound within the decision tree

The perturbation can have the effect of moving the design state to a different branch of the decision tree. The selection of the necessary perturbation is a difficult task based on identifying the single (or even multiple) modification to the topology likely to improve the overall design result. If the analysis of the perturbation requirement was correct then this new topology state should lie on the path of the optimal solution. The art of perturbation analysis can therefore be summarised as determining the network state to which to revert in order to reach the optimal solution (or optimal within an acceptable bound). Ideally this should be within the shortest possible number of steps since a perturbation that requires great computational effort may indicate that the original design process is poor. Evidently there are a myriad of possible perturbations, as there are also a myriad of additional design choices. The entire issue of perturbation selection is complicated by the change of route allocations following the removal of each link. All traffic demands using any removed trunks must be dropped. The status of the network therefore changes in two ways, the cost is lowered due to the reduced trunks but the traffic carried also drops across many other trunks due to the severing of any allocated paths.

With reference to the three-dimensional topology state tree it may be seen that a perturbation can lead to a complex change in the state of the network under design. The larger the perturbation the larger the disruption to established routing. It is concluded that it is important to minimise the size of the perturbation in order to prevent the destruction of a large portion of the traffic routing already established. It can also be deduced that in order to achieve design perturbation, it is possible to change the routing of traffic demands rather than the physical trunk allocations. For instance, paths established during an early phase of the design process may be beneficially rerouted at a later point (or indeed after the first design phase is complete) in order to find lower hop alternatives. This results in a lower overall trunk capacity requirement and the possibility for lowering total trunk costs. At the extreme it might be possible to reroute all the paths carried by a selected trunk, in order that it may then be deleted.

The effect of the perturbation on the overall design can be illustrated by the example in figure 5.7 below. The perturbation effectively allows the design to revert to a lower cost state by the deletion of capacity perceived to be inefficient. The topology may then be redesigned, preferably in a new direction to reach an overall improvement in the final solution, though there is no basis for assuming this to be the case in every instance.

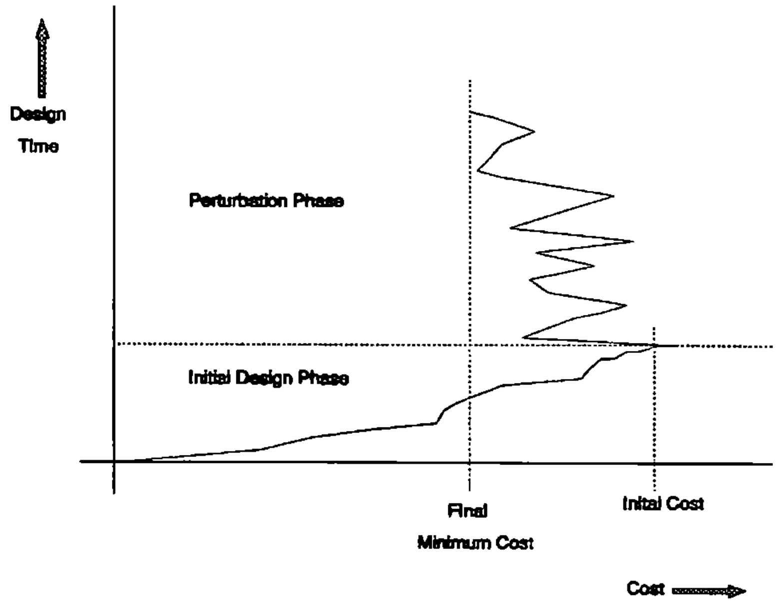


fig 5.7 Perturbation reverts to a previous topology

The quality of the perturbation analysis is critical to ensuring that the selected perturbation and subsequent forward progression may reach an improved topology. It is more than possible that multiple design and perturbation stages may result in a fluctuation of the overall cost which may be repeated until an acceptable minimum is reached or a prior one cannot be bettered.

5.7. Ordering Selections within a Threaded Design Process

In constructing the basic threaded design routine the cost of allocating each possible path to all possible traffic loads and then allocating the preferred choice it is clear that a number of methods are possible. For example a simple route allocation method might be of the form:

i. for all loads in the demand matrix

```
begin
find cheapest path for load
allocate cheapest load
end;
```

However, the simple selection process is weak since there is a random distribution of the traffic loads in the matrix if it is read out methodically.

It might seem reasonable to order the loads by their size before allocation, e.g. allocating the cheapest paths to the largest loads first. This would have the effect of establishing the largest volume of capacity in the network topology at the earliest moment, within the cost bounds of optimality offered by a greedy algorithm, thereby creating a topology dominated by the largest loads. However, there are considered to be two problems with this. Most importantly, taking the size of the load alone pays no attention to the location of the load's end points, if a situation is to arise where a large load exists across a long distance relative to the span of the network then there is no opportunity for nodes between the end points to tap into the high capacity trunk and benefit from the lower unit capacity costs. This is illustrated below:

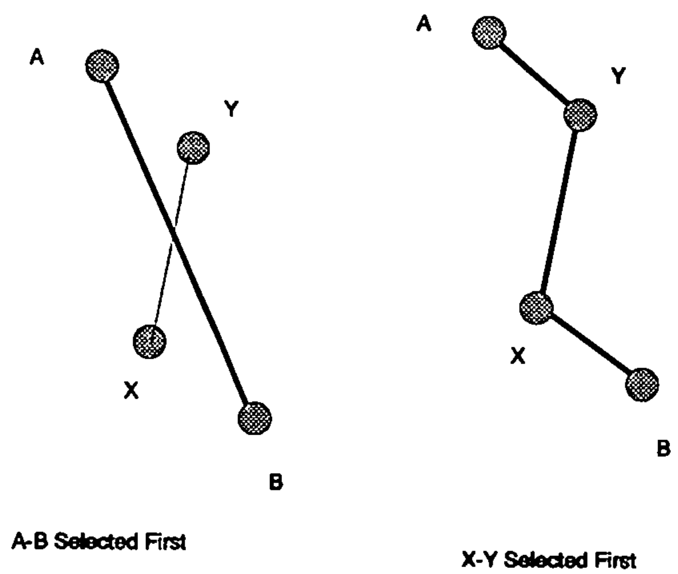


fig 5.8 Allocating largest demands has problems if trunks are long

If the largest demand, say between A and B, were allocated a direct path early in the design cycle then a smaller demand between X and Y would require a separate path. However, if the most economic paths are allocated in order then it would be possible for the A-B and X-Y paths to share a common high capacity trunk, reducing the overall cost if the increased path length is compensated by reduced channel costs of the larger capacity trunk.

The second problem occurs where there are a number of demands for the same number of channels, a further arbitrary scheme is then required to select an allocation order between them. Creating a rational rule for the additional ordering has not been found possible, nor would one offer any obvious design merit.

It was decided to use a greedy route selection rule where all possible (within a given bound, to be described later) routes are evaluated and the one with the lowest cost allocated. The cheapest routes will be those over shortest distances with the highest channel requirements since the cost function is proportional to distance and inversely proportional to channel size. As the cheapest routes are established other loads will be able to connect into them to gain further advantage of the low channel costs. this Preventing the problem illustrated in figure 5.1. It is acknowledged that a perturbation scheme is required to minimise any capacity overheads incurred as a result of the 'shortsighted' nature of the greedy algorithm. Perturbation schemes will typically try to reduce the level of spare capacity on each trunk and even find reroute options to free up and delete trunks if possible.

The outline format of the decision loop is:

i. for all traffic demands not yet allocated to paths

    begin

    calculate the cost of carrying each load

    end;

ii. allocate a path to the lowest cost traffic demand found in step i;

iii. repeat from i. until all loads allocated.

This would perform the first pass at creating a functional topology but a perturbation scheme would then be required perhaps taking on the form:

i. scan all links in the current network model;

ii. delete all links carrying less than X% (e.g. 50%) load;

iii. reroute all paths severed in step ii.

In order to implement this type of system it is necessary to examine the nature and consequences of such threaded decision making process, i.e. those where the decisions at each step are dependent upon the preceding ones. The use of such threaded methods have a range of benefits and drawbacks which will be investigated following the presentation of the basic groundwork.

While retaining sight of this inherent problem with threaded, greedy design methods, a general method can initially be built for implementing a threaded search with a separate perturbation scheme to follow. The general method involves the repeated execution of a search for the next apparent best move, in the manner of a chess player.

This can best be shown with a simple pseudo-program;

```
BestRoute = null
REPEAT
        Find the 'Best' available route from (i,j)  ----(1)
        UNTIL Tested all available routes
IF Route (i,j) Better than BestRoute THEN
        BestRoute = Route (i,j)
Allocate BestRoute    ----(2)
UNTIL all routes allocated (for all i,j; D_ij <> 0)
```

fig 5.9 Pseudo-code for the 'Route Finder' shell

(1) is the Route Finder algorithm that looks for the least weighted cost of all routes possible between nodes i and j,

(2) is the Master Selector process, detailed below, that makes the choice between all available minimum routes.

Even though this describes the overall method there is no description of what is meant by best route from i to j or even what is the best route to be selected at each pass of the outer loop.

The general network design method is typified by Saksena [1989] and Yokoyama [1988] who performed an initial solution phase where no 'quality of design' parameters were measured, only a topological (link based), feasible solution was sought. An alternative method is to create random starting networks [Gerla, Monteiro & Pazos, 1989]. The initial topology creation phase is then followed by a perturbation/optimisation scheme that performs the primary function of cost minimisation. The emphasis of these methods is placed on the quality of the optimisation phase. A significant problem is that results can be localised to a single minimum cost design, unable to reach a better globally optimum costed design.

5.8. Allocating Link Capacity to Backup Routes

The allocation of link capacity to primary route traffic is straightforward requiring only the addition of the capacity required by each traffic demand. In the model used by the design system a variable exists for each link and stores the accumulated channel requirement.

Devising a system for the allocation of capacity to backup routes is a considerable task. It has been shown in chapter 4 section 2.4. that Agarwal [1987] developed an analysis of backup traffic requirements based on the maximum capacity required by each link to cater for the worst case link failure. It is important to stress that this analysis method can only calculate the backup capacity requirements for ***known*** links and is therefore only of use for analysing a completed design. To be applicable in a practical situation it might be possible to apply the link failure analysis to a completed design, analyse the backup requirements and use the results to perturbate the design towards a lower cost. Unfortunately such a process would be very slow given the time required to perform the link failure tests. A further problem would be the unknown quality of any results and potentially lengthy perturbation stages.

The link failure analysis is largely inappropriate for use in a threaded design system where links are successively added and the network is incomplete until the last route is allocated. It is equally inappropriate to allocate exactly the demanded bandwidth for each backup route since, this would double the required network capacity. As the Agarwal link failure analysis shows, links both lose and gain traffic due to other links failing so the duplication of capacity would be unnecessary. Without analysing the actual load changes following a link failure it is not possible to predict accurately the requirement.

It is necessary for a method to be developed for allocating sufficient yet not excessive backup link capacity required by each route. Two approximations were developed following an analysis of the way in which traffic may be rerouted after a link failure, assuming only single link failures are likely. The first is based on historical data indicating 'likely' backup capacity requirements and the second is based on a 'Max Hold' system to track the greatest backup capacity requirement on each trunk. The need to approximate the backup channel requirements results in some variation between designs in the number of allocated backup channel. For this reason a measure of the ratio of primary to backup channel allocations are calculated. If the ratio is 1.00 there are equal numbers of primary and backup channels allocated. This does not imply a wasteful allocation (e.g. doubling) of capacity for backup paths since they will always have an equal or greater number of hops than the primary paths.

It is necessary to judge the number of primary and backup channels allocated by differing designs with a view to selecting the cheapest topology where the allocation is a minimum and the spare capacity a maximum. The volume of capacity allocated to backup provision is a function of the mean backup hops and the absolute number of channels allocated:

backup capacity ratio = backup channels / number of backup paths * mean hops.

5.8.1. Proportional Aggregate Backup Route Requirement

Perhaps one of the simplest methods for approximating the backup route capacity requirements is to analyse a number of completed (or even merely feasible) network designs and deduce characteristic figures. This requires that completed designs have each link failed in turn and the resultant backup channel requirements for each link are calculated. It is then possible to calculate a characteristic ratio of primary to backup channel requirements. Tests conducted with a simulation system (Witness, supplied by AT&T ISTEL Visual Interactive Services) showed that for typical networks with between 20 and 40 nodes a backup capacity requirement is of the order of 86% of the primary capacity requirement.

The accuracy of such a technique will depend upon how representative of likely future designs were the original networks used to calculate the primary to backup ratios. This places a heavy emphasis on using traffic demands with similar magnitudes and distributions, otherwise there would be the risk that dissimilar figures would greatly distort the backup requirements. For instance, if a small number of very large demands were present, and the network were sparely linked the backup requirements would be large. This may be compared to a heavily connected network with many small demands, where it would be possible to find many routes around a single link failure and the corresponding backup channel requirements would be very much lower.

A further problem with this technique is that the channel requirements are known only as integer multiples of 64kbps. Should the backup requirement factor be 71% of primary demand for example, then the integer storage of low channel requirements is either not possible or requires further approximation. For example, a 2 channel primary would require 0.71 x 2 channels for backup. This is 1.42 channels, yet can only be represented as an integer, so either one or two channels would need to be allocated, either of which would be inaccurate and lead to an over or underestimation.

5.8.2. Maximum Single Backup Route Requirement

The second backup route capacity allocation scheme developed was based on allocating just enough capacity to carry the largest backup demand on each link, known as the 'Max Hold' backup allocation technique. This was implemented by storing the maximum capacity requirement on each link in a backup route each time one was allocated. If a subsequent backup requirement was greater than the previously stored capacity then the new value replaced the original.

The advantage of this system is that it is simple to implement in software and, while being quite a coarse approximation, in practice it proves no less useful than the more complex system of percentage aggregate allocation. For a maximum backup capacity ratio of 1.0 there would be dedicated capacity to all backup paths. However, if a number of backup channels share the backup capacity, using the 'Max Hold' allocation method then this ratio falls. It is possible to lower this ratio since resilience to any single link failure does not require the duplication of all backup paths. However, in lowering the ratio too far insufficient capacity might be left to cater for the most critical link failures, i.e. a total of 50 backup channels allocated between 20 backup paths of mean length 2.5 hops represents a backup capacity ratio of 1.0, whereas taking the first entry in the table, 54 channels are allocated to the backup with a mean of 2.63 hops per backup channel for 34 paths in total, the backup capacity ratio therefore = 54 / 34 x 2.63 = 0.60

This shows that the 'Max Hold' method allocates approximately 60% of the capacity required to duplicate all paths for this example topology. The approximation is a result of using the mean backup path length.

5.9. Perturbation Schemes Following Greedy Optimisation

The approach of many researchers in the network optimisation field [Boorstyn & Frank 1977], [Frank & Chou 1972], [Frank, Frisch, van Slyke & Chou 1971], [Gerla, Monterio & Pazos 1989], [Saksena 1989], [Tobagi, Gerla, Peebles & Manning 1978] is to create initial solutions using LP methods and reduce the topology cost by applying heuristic perturbation schemes.

This is the method adopted in this research project, though replacing the LP initial solution with a threaded design process. It is possible to apply more than one reduction method, using various combinations of each until no improvement is found.

Two methods were developed for evaluating the possible reduction potential of each network topology. Both methods are considerably different from those more typically found in other network design systems. It is usual to implement perturbation schemes that either delete links or adjust link capacities and then perform a further complete analysis phase to ensure that all routes can be satisfied by the new topology. It is the method by which links are selected for deletion or addition that differentiates the various schemes. They are typified by the Cut-Saturation [Boorstyn & Frank 1977],[Frank & Chou 1972], [Gerla & Kleinrock 1977], [Pierre & Hoang 1990] and the Branch Exchange [Boorstyn & Frank 1977], [Gerla & Kleinrock 1977] algorithms.

The Branch Exchange method does not appear particularly effective since Gerla and Kleinrock state that it performs a search for local optima using all possible single link elimination and addition combinations which becomes very time consuming for more than 20 nodes. They also indicate more links may be exchanged though the execution times would suffer considerably.

The Cut-Saturation method is considered an extension to the Branch exchange method by using means to determines the 'best' links to exchange on the basis of their utilisation and cost, hence eliminating the need to perform exhaustive searches.

In an admission of the complexity of finding even acceptable perturbation schemes Gerla and Kleinrock [1977] state that for complex design requirements such as 'm' hop and 2-connectivity (as required in this project) 'the network design can be greatly enhanced by using man-computer interaction'. They note this has the benefit of allowing the network designer to observe 'and eventually correct the topological transformations of the computer'. It is interesting to note that no indication is given of how the human designer is supposed to identify the errors or weaknesses in the resultant designs.

5.10. Link Costs Influence Topology

The cost function and the way apparent link costs change throughout the threaded process have a considerable influence upon the network topology.

To define the cost of a route, firstly, let a route vector R be defined as a series of L nodes forming a single path such that

$$R = \{ r_1, r_2, r_3, \ . \ . \ . \ r_L \}.$$

It is necessary to use a mixed notation for the routes depending upon the particular instance and also define the links in the path as R_x, where x denotes the x^{th} link from node [x] to [x+1].

Restating the link cost equation from chapter 4, let c_{ij} be the unit channel cost between nodes i and j;

$$c_{ij} = l_{ij} \ [\ k_0(\Omega_{ij}) + d_{ij} \ \Omega_{ij} \, f(\Omega_{ij}) \ ] \ ;$$

$$where \ \Omega_{ij} = (\alpha_{ij} + \beta_{ij}) \ .$$

where	d_{ij}	is the distance from node i to node j;
	l_{ij}	is a [0,1] matrix representing the presence of link i-j;
	Ω_{ij}	is the allocated traffic variable representing the sum of primary and backup channels allocated;
	α_{ij}	represents the primary traffic on link ij;
	β_{ij}	represents the backup traffic on link ij;
	k_0	is the fixed link setup cost for the channel capacity Ω_{ij};

$f(\Omega_{ij})$ is the cost rate of change function, it determines the change in per channel cost with increasing channel allocation.

It is now possible to write the absolute cost of a route as being the sum of all included link costs, from r_1 to r_L. This assumes that no other connections are currently made in the network. Let the i^{th} link of the route be components that represent the link pair of the route,

$$c(R) = \sum_{i=1}^{L-1} [\, k_0\,(\Omega_{r_i r_{i+1}}) + d_{r_i r_{i+1}}\, \Omega_{r_i r_{i+1}} \cdot f\,(\Omega_{r_i r_{i+1}})] .$$

The total network cost at any given stage of the optimisation, step n, can be written by replacing all static variables with dynamic variables and a (n) time indices;

$$\psi(n) = \sum_{i=2}^{N} \sum_{j=1}^{i-1} l_{ij}(n) .\, [\, k_0(\Omega_{ij}) + d_{ij}\Omega_{ij}(n)\, f(\Omega_{ij}(n))\,] ;$$

$$where \quad \Omega_{ij}(n) = [\, A_{ij}(n) + B_{ij}(n)\,] .$$

$$Let \;\; K_0 = \sum_{i=2}^{N} \sum_{j=1}^{i-1} l_{ij}\, k_0(\Omega_{ij}) ,$$

hence

$$\psi(n) = K_0 + \sum_{i=2}^{N} \sum_{j=1}^{i-1} d_{ij}\Omega_{ij}(n)\, f(\Omega_{ij}(n)) ,$$

where the $l_{ij}(n)$ is a 0,1 matrix to represent whether a link from node i to node j is in place at stage n of the design process.

A UK cost analysis [Tarifica, May 1993] shows that the typical costs for digital communications links are as in the table below [BT prices].

Capacity	64kbps	128kbps	256kbps	512kbps	1Mbps	2Mbps
Setup Cost (£)	760	4480	4960	5920	7840	12400
Monthly Rental (£) (per km)	0.56	18.00	36.00	72.00	144.00	291.00

Table 5.1 Current BT trunk costs

Such cost functions can have important implications to the path selection procedure due to the way in which the cost reduces as the capacity increases. This occurs for two reasons, each of which is illustrated separately below.

In the figure 5.10 below one can see the effect of a constant cost per unit distance for all capacities with a setup cost that decreases with increasing capacity (VX axis). The cost reduction seen with increasing trunk capacity is therefore purely a function of the reduced setup cost, this effect will, as a result, be dominant for short trunks. The line AB cuts the link cost surface at a level of constant cost in the plane XVD. The line WY projected onto the lower XVD plane represents the distance a link of each capacity may cover for a constant cost. It may therefore be seen that as the capacity increases a link may be made longer for the same cost. This effect is found to be highly significant to the running of a threaded design process, because it implies that short links may be replaced by longer ones if they carry more traffic for no increase in cost. Alternatively, a route may be longer in distance and/or the number of links if it can utilise links of a higher capacity.

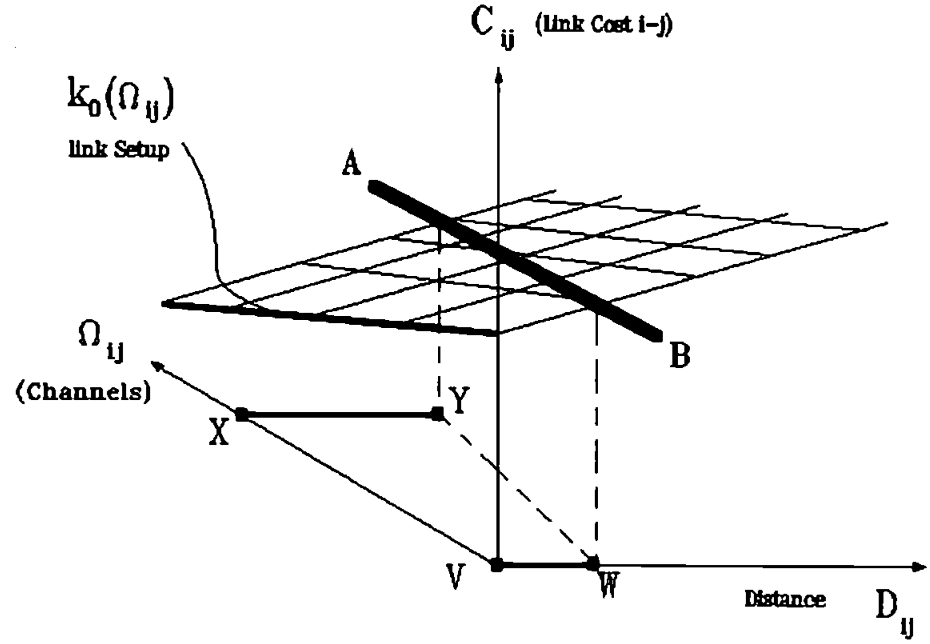


fig 5.10 The cost function - fixed cost per unit distance, falling setup cost per link for increasing capacity

A-B represents the constant cost level

V-W represents the distance possible for lowest channel allocation

X-Y represents the greater distance possible for highest channel allocation.

?

In figure 5.11 below, a similar effect to that above is observed, yet it is now as a result of a fixed trunk setup cost and a decreasing cost per unit distance with increased trunk capacity. This cost reduction effect is therefore going to be most noticeable on long trunks and negligible on short trunks.

The line AB is drawn at the constant cost level in the XVD plane. Projected onto the base of the graph, the line WY illustrates the increasing link length possible for equal cost as the capacity increases. In this example the distance increase is due to the reduction in overall link cost as a result of the falling distance related charge. The gradient in the CVD plane decreases with increasing capacity.

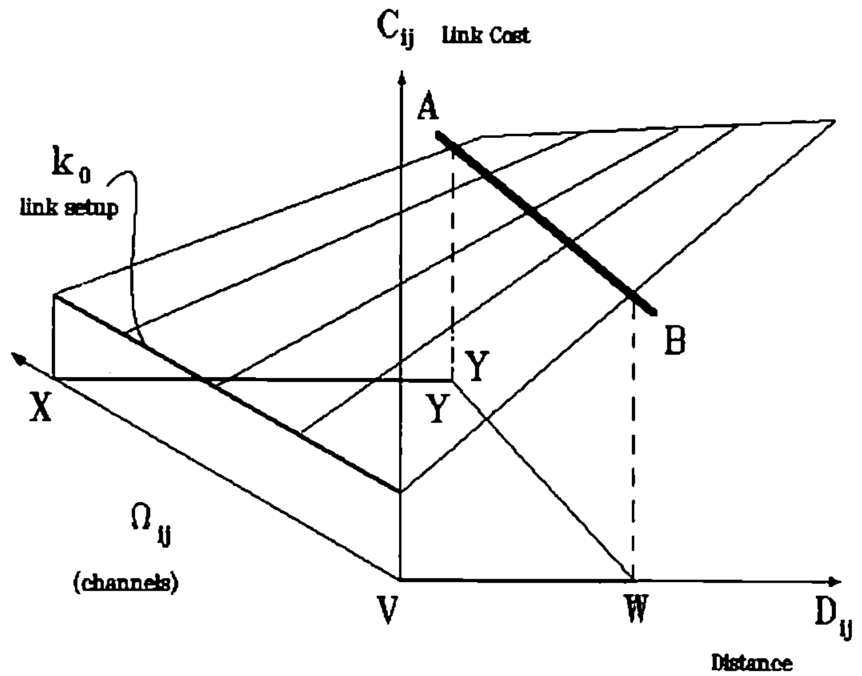


fig 5.11 Cost as a function of capacity and distance - cost per unit distance reduces with increasing capacity with constant setup cost

In practice the two effects shown above work in concert to produce significant cost reductions per kbps. This translates into a much increased distance a link or route may travel on a higher capacity trunk than a lower one for the same cost (assuming that incurred spare capacity is not charged).

A-B represents the constant cost level;

V-W represents the distance possible for lowest channel allocation;

X-Y represents the greater distance possible for highest channel allocation.

The implications of this cost reduction are not readily apparent until the selection of routes and links are examined in more detail. The selection between a number of candidate routes for allocation during the threaded design procedure is complicated by two further factors. Firstly, it is found that the reductions in cost of the various capacities occur suddenly at specific break points, and the true cost functions are not linear as shown but step functions. Secondly the very fact that the setup costs are applied for each link means that the more links there are in a route the greater the contribution of the setup charges to the overall cost and the higher the cost relative to a single link of the same length. The precise form that a trunk tariff might take in any selected country can therefore have a significant effect on the link cost tradeoffs used within the design process. In some countries it might be found that short trunks encourage capacity increases and in other countries long trunks might be favoured at higher speeds. The cost/ benefit tradeoff must therefore be applied separately on each route evaluation.

5.11. The Cost Improvement Measure

In order to identify the nature of the change in route costs as a network design progresses it is useful to look at the change in the cost of the network from step n to n+c of the optimisation. An expression is written to reveal the change in topology cost between one step, n and a future step at n+c for a network with N nodes.

$$\delta(n,n+c) = \psi(n+c) - \psi(n)$$

$$= [\ K_0(n+c) - K_0(n)\ ] +$$

$$[\sum_{i=2}^{N}\sum_{j=1}^{i-1} l_{ij}(n+c)\Omega_{ij}(n+c)d_{ij}f(\Omega_{ij}(n+c)) -$$

$$\sum_{i=2}^{N}\sum_{j=1}^{i-1} l_{ij}(n)\Omega_{ij}(n)d_{ij}f(\Omega_{ij}(n))]$$

where $\delta(n,n+c)$ represents the change in cost from step n to step n+c of the design process;

$\psi(n)$ represents the cost of the network at step n of the design process;

$K_0(n)$ is the sum of all link setup costs at step n of the design process;

$l_{ij}(n)$ is a 0,1 matrix to indicate presence of link between i and j at step n of the design process;

$\Omega_{ij}(n)$ is the sum of primary and backup channels allocated at step n of the design process;

d_{ij} is the distance between i and j;

$f(\Omega_{ij}())$ is the capacity related cost multiplier of the link.

This expression shows the difference to be in two distinct parts, firstly the difference in setup cost due to the change in number of links added. The second part of the cost difference is due to the change in link capacities.

It can be seen that the network costs can be limited by two means:

i. the number of links allocated can be kept minimum to keep the K_0 figure low;

ii. the traffic carried on each link can be maximised to gain from the cost improvements due to the negative gradient cost function.

5.12. Deferred Cost Benefits in Threaded Processes

In the design of a large network there might be many routes requiring selection but the benefits of reduced overall cost from the use of the maximum hop selection within the desensitivity bound may not be immediately gained, if at all. To understand how this might occur, consider the problem of the time sequence of route selection.

By increasing the traffic on a link at step p of the design process then the per-channel cost of that link on subsequent steps p+q will on average be lower due to the decreasing unit channel cost function.

In practice, the improvements will not be experienced immediately, but will appear when the link concerned is next used as part of a route and its contribution to the route cost is evaluated. This improvement comes at some later step $p + r$ where $r > 1$. The cost improvements must therefore be considered to affect future route selections with a probability dependent upon:

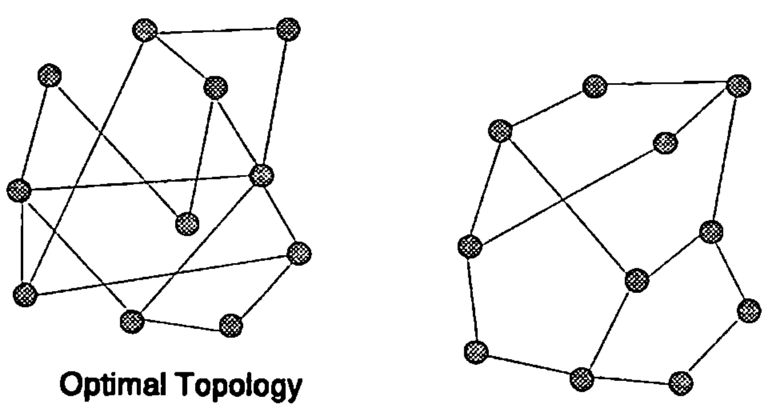
i. how many link demands have been allocated;

ii. how many link demands are left to allocate;

iii. the capacities assigned to each link - the higher a link's capacity the lower its per channel cost and hence the more attractive it is for use in other routes.

A further problem arises when the cost function is banded, not continuous, and in practical design problems this is by far the more common situation. This results in zero improvement if new capacity is added within the same cost band and hence there is no reduction in the unit channel cost.

It is now possible to envisage a point at which the mesh ceases to be cheaper than the ring by comparing the cost savings of direct routes (minimum distance) and the cost benefits of higher link capacity. The critical difference is determined by the nature of the cost function. In extreme terms, if the setup costs of links were dominant then the minimum link network, the ring would prevail. If the cost were heavily dependent upon the capacity carried and largely insensitive to distance then the full mesh network would result. If the distance related cost were dominant then the network would combine attributes of the ring and mesh dependent upon the particular traffic demands.



Optimal Topology
for Dominant Capacity Cost

Optimal Topology
Dominant Link Setup Cost

fig 5.12 Dominant costs influence topology

As indicated before the increase in overall channel capacity due to the selection of longer hop routes can lead to a critical minima where excess may cause the global network cost to rise. It is foreseeable that some 'overshoot' in the allocation of capacity might occur during the design process. This will be due to the need to balance the addition of capacity in the problem just illustrated above, where the allocation at stage n only yields savings in the future at stage x+y. Excessive allocation might occur as the network design phase comes to an end or on links that are allocated capacity early in the design and not adjusted later.

This leads to an inevitable likelihood of some excess capacity being allocated and it is therefore seen that some form of final perturbation in the design method is required. The final capacity allocations may therefore be usefully analyzed as part of a perturbation scheme to adjust any routes that can be more economically rerouted, perhaps via fewer hops, over links that were allocated subsequent to their original allocation. This may effectively remove unnecessary additional capacity and offer a significant design benefit.

The issue of allocating greater hop routes in the expectation of future improvement in cost function is complicated further by the problem of over allocation of capacity. Since a further reduction stage is required the design process cannot therefore be considered complete when all routes are allocated. The over allocation of capacity may lead to overall cost improvements when simple redundancy tests are performed, rather than if a more conservative allocation measure is used. This is because the network topology will consist of a mixture of large capacity links with lower channel costs and a few low capacity links. It is preferable to delete the more costly (by unit channel cost) low capacity links and reroute the displaced traffic over more economic links. This will lead to lower overall design costs, making use of later trunk allocations for those early route selections.

5.13. Discontinuous Cost Functions

It is found that cost functions implemented by network capacity providers are not smooth continuous functions and this has an effect on the cost analysis performed above.

What has been seen up to this point, is an indication that there is potential for network cost minimisation by selecting the route with greater number of hops if costs are similar. The range within which we consider routes to be of similar cost must therefore be carefully defined. The potential savings in the future design must not be less than the cost difference of the two routes being compared for selection.

The implication of the real world cost function is that line capacities within each band can be viewed as having two separate cost measures, the choice being made by the designer. If a link requires a given channel capacity then the minimum number of channels allocated must be sufficient to satisfy all those demanded. Due to the banding scheme there are likely to be excess channels available.

There are two charging viewpoints with regard to spare capacity:

i. the cost of the spare channels is equally divided between the channels allocated -this results in increased per channel costs to the user and therefore discourages the allocation of higher capacity links;

ii. the cost of the spare channels is absorbed by the network designer and effectively paid for by the network provider company. This is the most beneficial system to optimisation because it encourages the allocation of high capacities and the lowering of per channel costs. The network provider can view the extra cost in two ways;

   a. the cost should not be excessive otherwise profitability is affected, it is up to financial controllers to decide what is an acceptable level of cost excess;

   b. the cost should not be a major concern because networks are assumed to be both dynamic and expanding, the excess capacity will soon be soaked up by expansion demand.

CHAPTER 6

A New Network Design Method

6.0. Introduction

This chapter details the specific problems found in developing a traffic based design system and looks at a number of heuristic methods developed to reduce the computational complexity of finding 'good' solutions. A new insight into the large scale network design problem is gained from this research and the major influences it has on the new design method are detailed. Much of the development work was aided by the testing of algorithms on network design scenarios. Analysis of the relative quality of the results for the various methods tested served to highlight those important areas of the problem requiring special attention. A sophisticated 'window and mouse' based software shell was written to allow experimentation with many different design techniques. The 'shell' greatly speeded development and testing of designs by offering a ready-made testbed for the specification of node positions, traffic requirements and routing parameters with full interrogation of the results being possible when designs were complete. Full support for the saving and retrieval of all data and parameters was provided. The analysis of network designs was facilitated by viewing routing tables, lists of routes using each link and analysing the network costs apportioned to each route allocated. These functions are made available in separate windows of the design software and all results may be printed out in report or graphical form.

All network design scenarios are based on circuit switched core networks and require the permanent allocation of bandwidth between each circuit source and termination. This is consistent with the sharing of network resources between many customers and various network-wide services on a large scale international network. It is assumed that backup paths are automatically selected should any link in the primary path fail.

6.1. Summarising the New Network Design System

In order to produce a new design method a range of ideas were investigated and the chosen method improves upon the typical greedy algorithm's 'select instantaneous best' rationale. The general principle is to focus on the main short-term gain problem found with greedy methods and move away from the notion that the 'best' selection is required at each decision point. The assumption that the final outcome of the design is not known may be considered pessimistic, it is possible to develop an understanding of what topological features constitute likely components of good designs. An analysis of the trunk cost structure and backup capacity requirements has shown that a tendency to create loops may minimize total cost where trunk setup charges dominate, as is the case between countries in Europe. This information can be used as a second criteria, where cost is the first, in a design method. It is proposed that a better general principle is to make a primary selection between best route selection on the basis of cost but decide between similar cost routes on the basis of their likeliness of being in a good final design. The objective ~~being~~ is to steer the design towards having those characteristics known to be desirable, for example hop count and nodal connectivity. The achievement of this 'steering' is investigated further.

In summary the design system is based upon a systematic search for the best route to satisfy each traffic demand. The best route is selected primarily on the basis of cost but, where the cost difference between competing routes is small, the selection may beneficially be made on a secondary criteria with a more global view of what constitutes the best route. Since the number of possible routes to be searched for in any practically sized network is extremely large, a method is also required to speed the search process. This may neglect to search for routes in outlying parts of the topology without overlooking the most likely routes to form an optimal design. In order that any redundant capacity or non-optimal routing in the resultant topology may be removed the initial design phase is followed by a design perturbation phase.

The new method is intended to maximise the allocation of routes likely to be components of a good, completed network design. Such a process is expected to have a higher probability of producing a final design within a narrow bound of the optimum solution than one that merely uses a shortsighted 'select immediate best' greedy method which can produce results of unknown quality. A further aim is to produce a means by which alternative topologies can be designed for the same set of input traffic requirements. This requires that the design algorithm has a 'tunable' parameter, which is preferable to randomly adjusting input data to stimulate different topologies. The benefit is that a number of slightly different designs may be produced which will all exactly meet the same design requirements. This then frees the network designer to exercise some free choice on the basis of political or strategic factors in the knowledge that the selection is not compromising the input parameters. Since the network design problem is known to be hard and heuristic methods are used to limit the design time it can reasonably expected that any tunable parameter does not have a single optimal value, but rather, requires some measure of adjustment for a given traffic, trunk cost and node configuration.

This chapter details an analysis that attempts to judge the quality of routes based on the expectation of their inclusion in a final solution, i.e. how likely is it that a given route will be a component of the optimal solution. A prerequisite of this is an understanding of what constitutes a 'good' network design. To simply state that a 'good' design is of low cost ignores the importance of the inherent design characteristics. Such characteristics can only be seen from analysis and may be expected to be of the form of the average number of links per route, both for primary and backup, the average link utilisation and the selected capacity. The way in which spare capacity on links is distributed might prove to be important in allowing for traffic growth in favoured areas. Whatever the characteristics are, they may be used to influence the selection of routes and links in the design phase in order to bias the design in what are known to be favourable directions.

6.2. The Cost of Routes in a Threaded Design Process

It was shown in section 11 of chapter 5 how the apparent cost of links changes throughout a threaded design process. In order to determine how this may be used to benefit the route selection process it is necessary to first formulate an expression showing the difference in cost between two routes. The cost of a route is based on the sum of the pro-rata costs of all links in the route. Extending the link cost function from chapter 5, section 11 to create the route cost, c_r, for an L link path gives,

$$c_r = \sum_{i=1}^{L-1} [\, k_0 (\Omega_{r_i r_{i+1}}) + d_{r_i r_{i+1}} \Omega_{r_i r_{i+1}} \cdot f(\Omega_{r_i r_{i+1}})]$$

This represents the sum of all setup costs for each link plus the sum of all distance and capacity related costs. If the subscript notation is simplified by replacing the r,r+1 with i, thus

$$c_r = \sum_{i=1}^{L-1} [\, k_0 (\Omega^i) + d^i \Omega^i \,.\, f(\Omega^i) \,].$$

To simplify the equation further, replacing the sum of setup costs with K_0, where

$$K_0^L = \sum_{i=1}^{L} k_0 (\Omega^i),$$

to give

$$c_r = K_0 + \sum_{i=1}^{L-1} d^i \Omega^i f(\Omega^i).$$

By comparing two routes of differing length L_1 and L_2, so that if they are within a cost bound Δ of each other, then they may be evaluated on the basis of their number of hops. An expression may be written to test for the costs being within the bound,

$$| \sum_{i=1}^{L_1} (K_0^{L_1} + d^i \Omega^i f(\Omega^i) \pm \sum_{j=1}^{L_2} (K_0^{L_2} + d^j \Omega^j f(\Omega^j) | \leq \Delta$$

where Δ is called the desensitivity factor. The desensitivity factor is so named because it represents a cost boundary within which two routes may be evaluated on the basis of a secondary criteria on, the route selection decision is therefore desensitised to the absolute cost of routes by this factor. The final expression for c_n above forms the basis of the route selection process. It is now necessary to determine its precise application.

Clearly the cost of various route options are compared many times during the execution of the design program. The general method is to select the cheapest route at each stage. If the cost of a new route is within a certain range (the cost desensitivity bound) of that for the 'best' route then the selection will take place based on a secondary criterion. If the cost of the proposed route is lower than that of the running 'best' by more than the desensitivity range, then the new one will be selected to become the new 'best'. This provides a means for valuing the secondary criteria through the use of the desensitivity range while not wasting time performing extra evaluations when the new cost is clearly better or worse than the running best. This has a great effect on the route search time since the simple cost comparison is faster than that based on the combined cost and secondary criteria comparison.

In looking at this equation it can be seen there are a number of influences on the route and hence network cost:

i. reducing the per channel cost will bring down the overall route costs;

ii. as the route takes a longer path the cost rises;

iii. if two routes have approximately the same path length and one uses more hops then the other then the one with more hops will contribute most to raising the channel allocation and hence reducing the per channel cost;

iv. if the secondary objective is to increase the number of hops in routes it must only be done within a limit where any increased path distance does not raise the cost to the detriment of the overall topology.

Hence it is possible to see that selecting the greatest hop routes, within defined cost limits, will contribute to raising channel allocations and hence lead to a lowered unit cost. It is essential that the increase in channel allocations is sufficient to benefit from the unit channel reductions but not excessive in a way that additional nett cost is incurred.

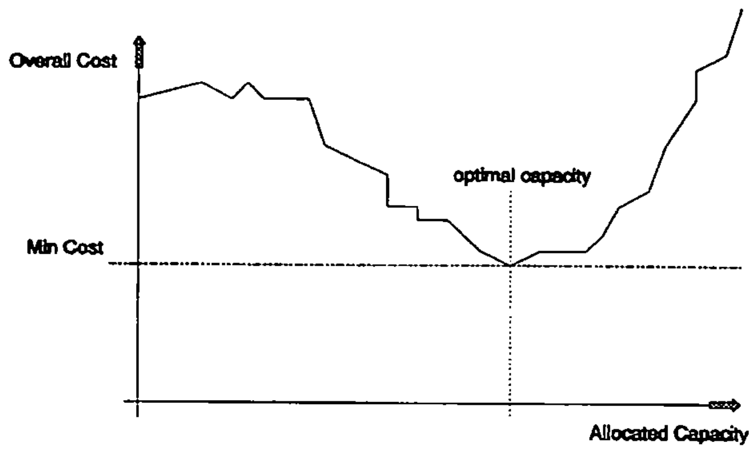


fig 6.1 Diagrammatic relationship between the number of channels allocated and network cost

The diagram above illustrates how, initially, a decrease in network cost results from an increasing allocation of capacity. This is due to the reduced unit channel cost of higher capacity trunks. A rise in cost then results when the critical minimum point is exceeded and the increase in allocated capacity no longer contributes to more effective channel utilisation but is merely an unproductive overhead.

6.3. Route Cost vs Length

Having written an inequality expression, in section 2, describing the limiting difference in cost acceptable for two given routes, it is possible to explore the manner in which route costs may differ within the desensitivity range.

As the allocated traffic and hence capacity of links increases the unit channel cost falls, allowing routes to travel greater distances on high capacity links than for low capacity links at the same total cost. This analysis is distinct from the hop-based route analysis which merely looks at link connections between nodes and not route connections. For a single hop route (direct from source to sink) the comparison of capacity discount against distance is simple, the excess distance a link might take at a higher capacity is proportional to the reduction in unit channel cost per unit distance.

For routes made up of two or more links the comparison of costs and distances becomes more complex due to the combinatorial problem of all possible link capacities on each route. However, an algebraic function can be written to describe the manner in which unit channel cost fall due to the trunk tariff structure, allowing routes of increased distance to be considered.

In a simple analysis of a three node scenario with nodes ABC in figure 6.2 below, it is necessary to consider the difference in cost between the direct route from A to B and the indirect route ACB:

i. if the capacity on links AC.BC.AB is equal then the decision is simply a shortest distance selection, the unit channel costs are equal and the route direct from A to B will be selected as being the cheapest;

ii. if there is disparity between the capacities allocated then the unit channel costs of the links will influence the decision. If there are larger channels allocated between AC and CB than the direct path AB then the unit channel costs will (usually) be lower. The selection is then based on the cost savings due to reduced channel cost and the additional cost of the increased path length.

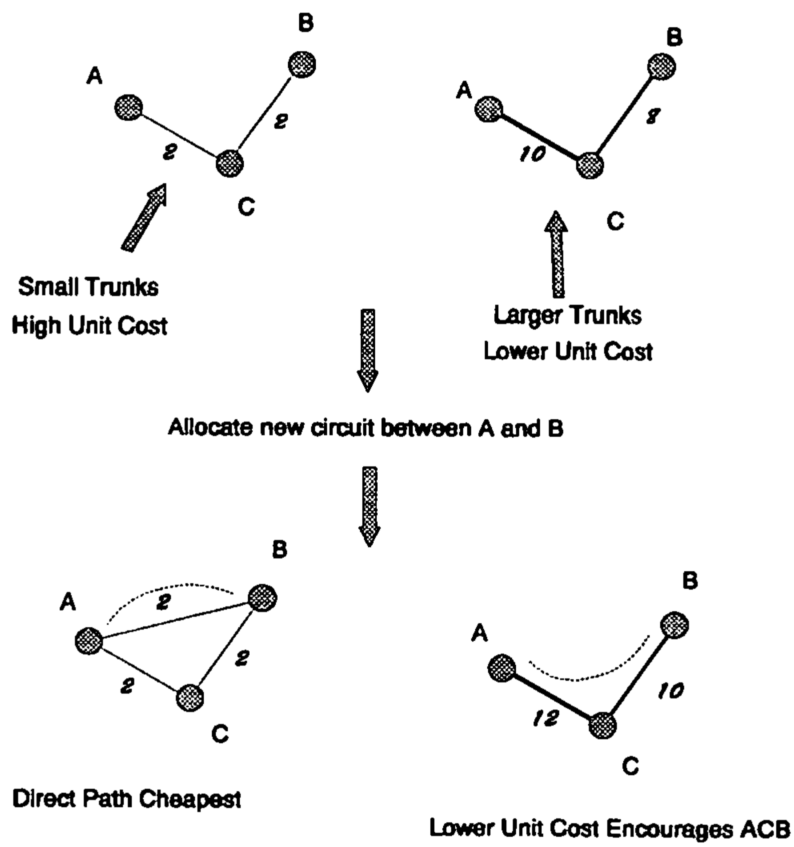


fig 6.2 Differing capacities in triangle influence decision

To simplify the initial analysis consider the links AB and AC to have zero capacity allocated. The situation where the capacity on link BC is greater than one channel implies that the unit channel cost of BC will be less than that of AB per unit distance. The increase in distance that link ACB may travel is due to this saving in unit channel cost. This may be analyzed by first defining the change in link i-j cost as $\delta_{ij}(D)$ where D is the change in allocated channels. Thus

$$\delta_{ij}(D) = k_0(\Omega_{ij} + D) + d_{ij}(\Omega_{ij}+D)\, f(\Omega_{ij}+D) - [\, k_0(\Omega_{ij}) + d_{ij}\, \Omega_{ij}\, f(\Omega_{ij})\,] .$$

In order that the bound on the excess distance via ACB be found, the cost of routes AB and ACB are equated, i.e. when the cost is the same, the limit of the boundary is defined. An expression is therefore written where the boundary is defined as the point at which the cost increment for routes allocated A-B equals that for A-C-B:

$$\delta_{AB}(D) = \delta_{AC}(D) + \delta_{CB}(D) \quad ;$$

which may be expanded to

$$k_0(\Omega_{AB}+D)-k_0(\Omega_{AB}) + d_{AB}\,[\,(\Omega_{AB} + D)\,f(\Omega_{AB} + D) - \Omega_{AB}\,f(\Omega_{AB})] =$$

$$k_0(\Omega_{AC}+D)-k_0(\Omega_{AC}) + d_{AC}\,[\,(\Omega_{AC} + D)\,f(\Omega_{AC} + D) - \Omega_{AC}\,f(\Omega_{AC})] +$$

$$k_0(\Omega_{CB}+D)-k_0(\Omega_{CB}) + d_{CB}\,[\,(\Omega_{CB} + D)\,f(\Omega_{CB} + D) - \Omega_{CB}\,f(\Omega_{CB})]\,.$$

The rearrangement places a constant term on the left hand side and leaves the right hand side setting the limit on the distances AC and CB, revealing an equation of the form:

$$K = d_{ij}\ v\ f\ (v) + d_{jk}\ w\ f\ (w)\ .$$

The right hand side now gives rise to a series of contours contingent upon d_{ij} and d_{jk} and the v and w values.

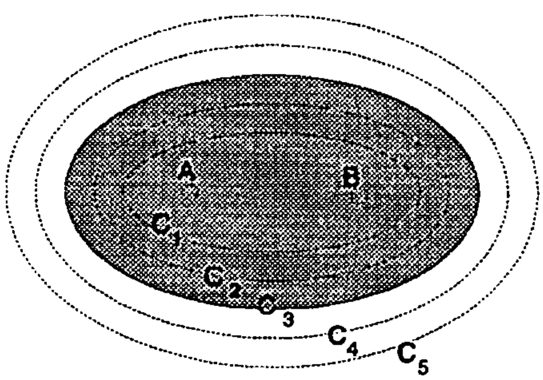


fig 6.3 Ellipses of constant inter-nodal distance

Figure 6.3 shows the range of possible positions node j may take, for each capacity C_1 to C_5, to be within the region to consider an alternate route. The boundary therefore represents the cost contour of each capacity. It is possible to examine any given network and evaluate the possible cost trade-offs of all possible routes, and hence the potential saving, by selecting routes of greater hop distance to gain overall cost savings from increased capacity and hence unit channel cost.

This may be extended to the variable case where the trunks AC, CB and AB may take on different capacity values and the boundary of the ellipse will change. For any pair of nodes with a given trunk capacity between, it is possible to calculate the maximum extra distance beyond the direct path length that the route may take. This takes into account the unit channel cost reductions gained by sharing other nearby trunks that are available. This then gives rise to the regions within which intermediate nodes may be used for routing. For a simple single intermediate node, assuming equal channel costs for the two hop alternative path, this is defined by:

$$X.c_1 \le Y.c_2;$$

where c_1 = unit channel cost of the direct path;

c_2 = unit channel cost of links between intermediate nodes;

X = distance between nodes A and C;

Y = total distance between A,B and C.

The extra cost of reaching intermediate nodes must be less than or equal to the saving made by the increased capacity allocations on those intermediate links. If the capacity required for the direct connection is c_1, the connection is 100km long and costs 100 units/km with an intermediate node already connected between A and C at a capacity c_2 with cost 50 units/km, the maximum distance ABC is given by:

$$100.100 \ge Y.50$$

$$\therefore \; Y \le \frac{100.100}{50}$$

$$\therefore \; Y \le 200.$$

Figure 6.4 below shows how, for equivalent cost, the distance to the intermediate node may increase as the capacity on the trunks rises and the channel costs drop.

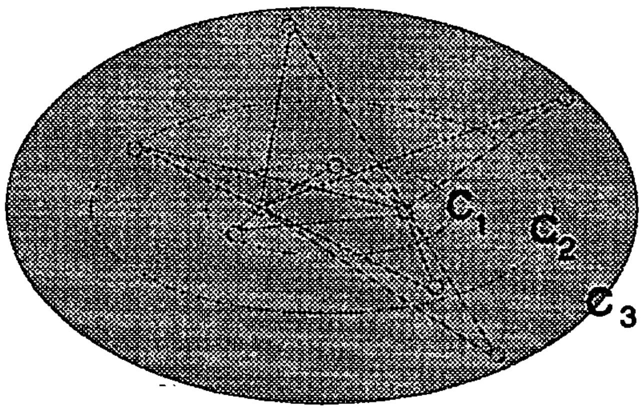


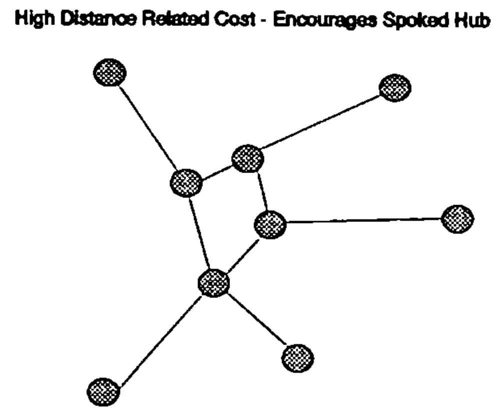
fig 6.4 Capacity increases lead to larger elliptic regions for alternate routes

6.4. The Influence of Trunk Cost Function upon Topology

The trunk cost function has an important effect upon the selection of routes, the number of hops per path can be as important as the physical path length. Section 3 has just shown that, within bounds, there are instances where a preference for selecting maximum hops routes will contribute to an overall lower network cost than by minimising distance or hops. The benefit is not immediate, but rather forms a long term strategy within the design cycle. The diagrams below illustrate topology features based on path hops.

The ring network, under certain cost constraints offers the minimum cost connection for N nodes. Where N is large the ring may no longer be optimal since there exist many hops per path, instead a network containing a number of large rings becomes optimal. For this to occur the cost function must show a significant reduction in the distance related link cost as the channel capacity rises. This then allows path lengths to increase as capacity allocation increases. In the extreme, where the distance related cost is negligible, the link setup cost dominates and hence the minimum cost topology will be based on the minimum number of links. No real network design problem is likely to suffer from any extremes in cost function or traffic requirements, but the basis for the formation of loops is clear.

In addition to the cost structures, the traffic demand structure significantly influences topology. In regions of a network where demand is sparse, connections to areas of more dense demand may be in the form of a 'spoked hub'. The diagram below shows how two different cost structures can persuade or dissuade the formation of such a feature.



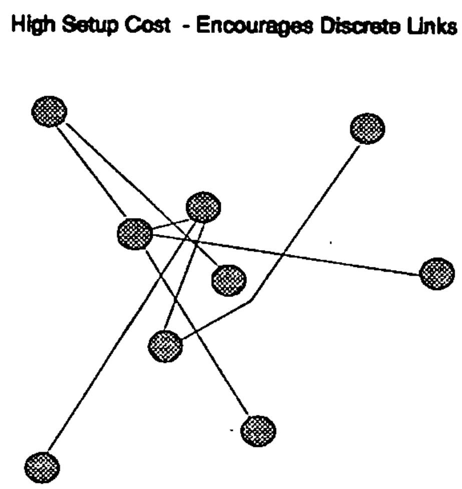


fig 6.5 Single links persuaded or dissuaded from forming a spoked hub

If resilience to single link failure is required then the effect is much the same but with the dualing of trunk access to the hub.

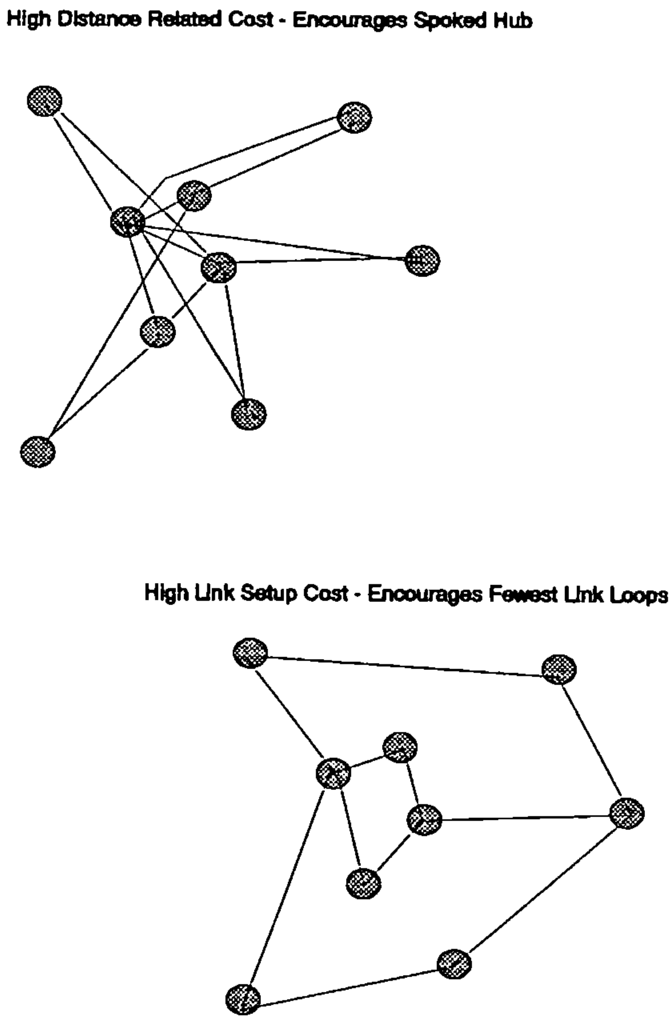


fig 6.6 Tariff structure can persuade or dissuade the resilient spoked hub

The essential principle of the link cost functions used in Europe and the USA is that the per channel cost reduces as the channel allocation rises, this function is discrete and concave.

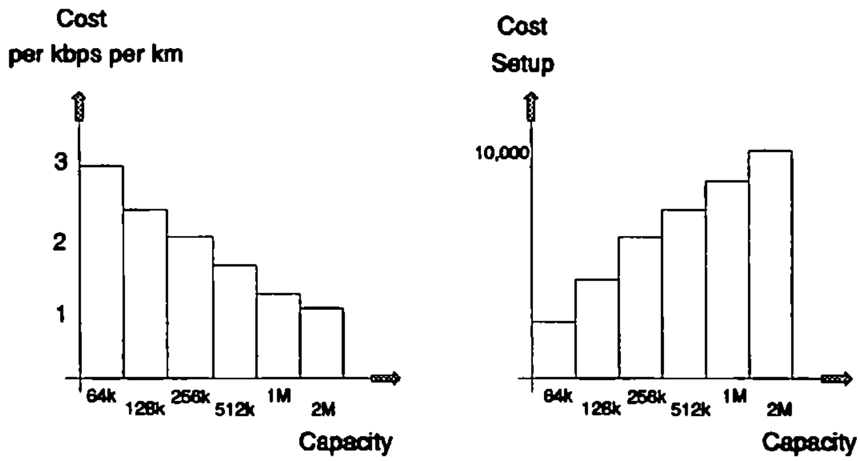


fig 6.7 Typical European tariff structure

From the point of view of the secondary route selection objective, a ring has the maximum total number of primary and backup hops between N nodes over N links and is a desirable component feature of larger networks. A resilient core network can be considered to be made up of a number of rings interlocked and overlayed. Instead of designing networks for minimum path length it is proposed that loops in a network design are encouraged by selecting between similar cost routes by choosing those with the most hops.

It is possible to take advantage of the observation that loops are potentially desirable. A route selection process may profitably select maximum hop routes without sacrificing the cost beyond a limit predetermined by the desensitivity parameter. It can now be seen how the desensitivity factor ties in with the route selection: routes are only selected on the basis of their number of hops if their total cost is within an acceptable range of the running best route. This method has the potential to sacrifice some immediate cost improvement for a future, overall network cost saving.

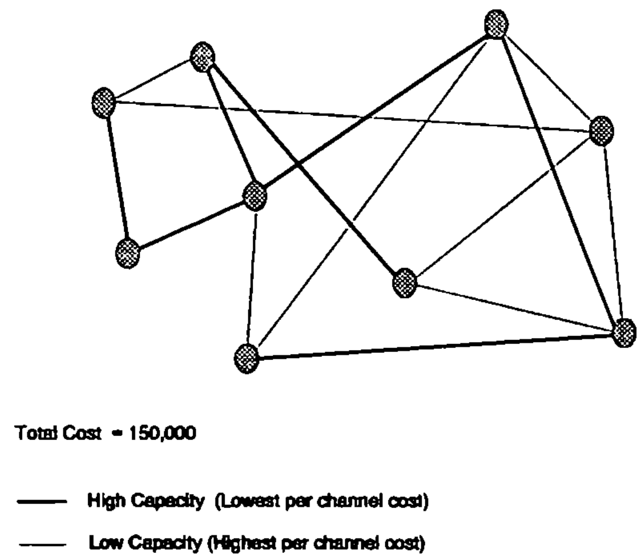


fig 6.8 A simple network topology design result

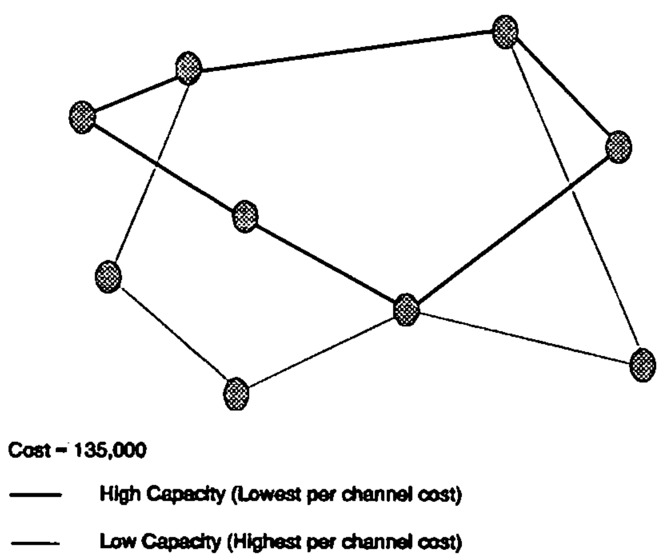


fig 6.9 Desensitivity methods encourage loops and reduce costs

In order that the route selection process be implemented to encourage rings, or loops in the network designs, the secondary route selection criteria minimises primary route hops and maximise backup route hops. This will only occur when costs are within the desensitivity bound and therefore it is not expected to lead to a rise in the overall network cost, rather to provide the bias towards lower unit cost and improved overall cost.

The route selection rule is therefore '*within a defined cost bound, select shortest hop primary routes and longest hop backup routes*'.

As stated previously there are two places in the design algorithm at which this rule needs to be applied; firstly the individual route selection for each node pair, and secondly the overall node pair route selection. i.e. for all routes between nodes i and j, for all nodes i and j.

However there is not enough evidence at this stage of the analysis to reveal whether the rules should be implemented in precisely the same form at each point. That is to ask 'is the shortest primary selected for each node pair, and then the longest, shortest primary selected' or may this rule be inverted and should the same be true for the backup route ? It is important to develop further the understanding of route costs and their impact on overall network cost before determining the use of these rules.

6.5. Implementing Cost Desensitivity in Route Selection

The selection between two routes takes place during the search through all possible routes between two nodes. A running best route is compared with all new routes found. The desensitivity, *d*, is used as a percentage factor of the best cost against which the new route cost is compared. This is illustrated below:

```
IF New Route Cost <(Best Route Cost x (1 + d))
AND New Route Cost > (1 - d) x Best Route Cost THEN
        Perform Secondary Selection (New Route, Best Route);
ELSE
        Retain Original Best Route.
```

fig 6.10 The cost desensitivity method

If the new route cost is less than the best cost minus the desensitivity cost then the new route is selected, essentially because the cost saving is so great, otherwise, if the cost falls within the desensitivity range the route selection is made on the basis of the secondary criteria. In the event of routes having an equal number of hops there is no penalty for selecting the cheapest.

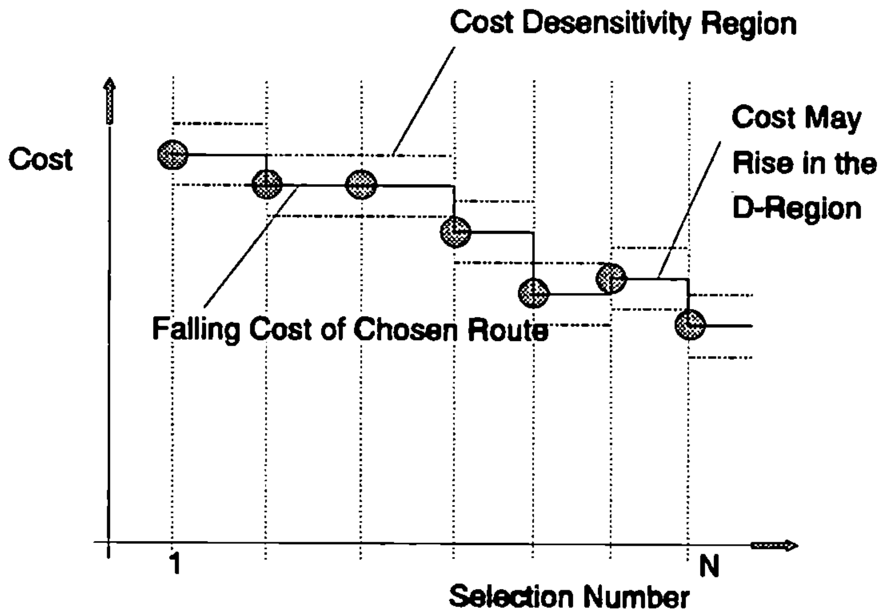


fig 6.11 Cost desensitivity region (D-region) relative to running best cost

The desensitivity selection algorithm in figure 6.10 is referred to as the select 'best' process and its use in the overall route selection algorithm is shown below in figure 6.12.

```
For all node pairs
        begin
        For all routes between node pair
                begin
                select (1) 'best' route
                end

        select (2) 'best' of the 'best' routes
        allocate best route
        end;
```

fig 6.12 Pseudo-code example of simple route selection

In the pseudo-code shown above in figure 6.12 there are two selection processes. The hop selection rule for each was closely examined following experimentation which showed that, of the four possible combinations of two rules for primary and backup route selection, there is only one that achieves the desired result of optimal loop creation.

For each of the primary and the backup routes it is possible for the outer selection (2) of figure 6.12 to choose the shortest primary and longest backup. Even though the routing selection is being applied to links that are not yet present but being 'designed' the rules for primary and backup are broadly the same as if the links existed. This means that the primary route will take the shortest path around the loop and the backup will take the longest.

The problem now is to achieve the main objective of maximising the number of hops in routes to gain the advantage of networks with loops. The aim is to maximise the minimum primary route length and maximise the maximum backup path length - each subject to the cost exchange limits. Referring to figure 6.12, this is achieved by making the inner route selection process (1) search for the minimum hop primary routes and maximin hop backup routes (always within the cost desensitivity bound). The outer selection (2) then finds the maximum hops of those routes from the first loop. The overall lowest cost route that meets the selection criteria is then allocated and the process searches for the next route. Eventually all routes are allocated as the outer loop searches for routes between all node pairs with traffic demands.

```
FOR all node pairs
        begin
        Route Finder:
        FOR all routes between node pair
                begin
                Route_Selection:
                IF ( COST (Route) < COST (Best Route) + Desensitivity) THEN
                        begin
                        select [ MIN_HOP (Primary Route) ]
                        select [ MAX__HOP (Backup Route) ]
                        end
                end

        Master_Selection:
        select [ MAX [ MIN_HOP (Primary Route) ] ]
        select [ MAX [ MAX_HOP (Backup Route) ] ]
        allocate Best Primary Route
        allocate Best Backup Route
        end;
```

fig 6.13 Pseudo-code example of Min-Max desensitivity route selection

The Route Finder acts as the initial filter by rejecting poor routes, but must make a selection of the best overall routes for nodes i-j. The best route is defined as the longest backup or the shortest primary. The outstanding issues to resolve are the setting of the desensitivity bound and the method of calculating route costs.

6.6. The Likelihood Route Costs within the Desensitivity Range

In order that the second objective of the cost desensitivity selection method be controlled it is necessary to consider the number of routes that fall within the desensitivity bound. With reference to the probability distribution function, pdf, of route costs below in figure 6.13. If the profile is peaked then many routes are of similar cost and a great many must be compared on the basis of cost and the second selection criteria.

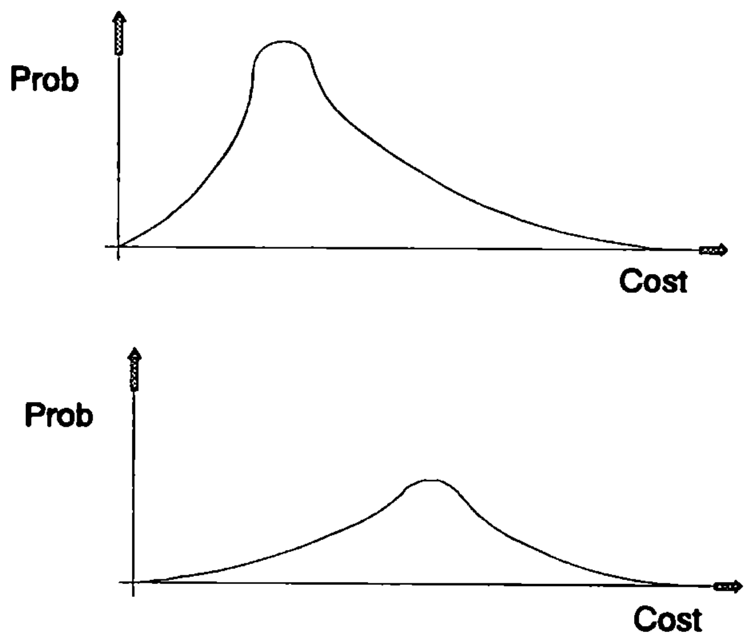


fig 6.14 Typical pdf for route cost between nodes i-j

It is therefore desirable that the number of routes of similar cost (within the desensitivity bound) are low to minimise the number of comparisons. It was found that implementing the desensitivity bound as a percentage factor of the running best route cost had the effect of making the selection very broad when the first routes of high cost were tested and progressively narrower as the 'best' route cost lowered. If the first route found was of a low cost it immediately removes the need to test a great many more expensive routes. The percentage value therefore acts as a data sensitive filter, narrowing quickly to the preferred lowest route costs. The range in route costs across the typical network is very large, typically of the order of 10:1. It is not necessary to make the desensitivity value small in order to provide rapid reduction in the number of route comparisons made on the second objective.

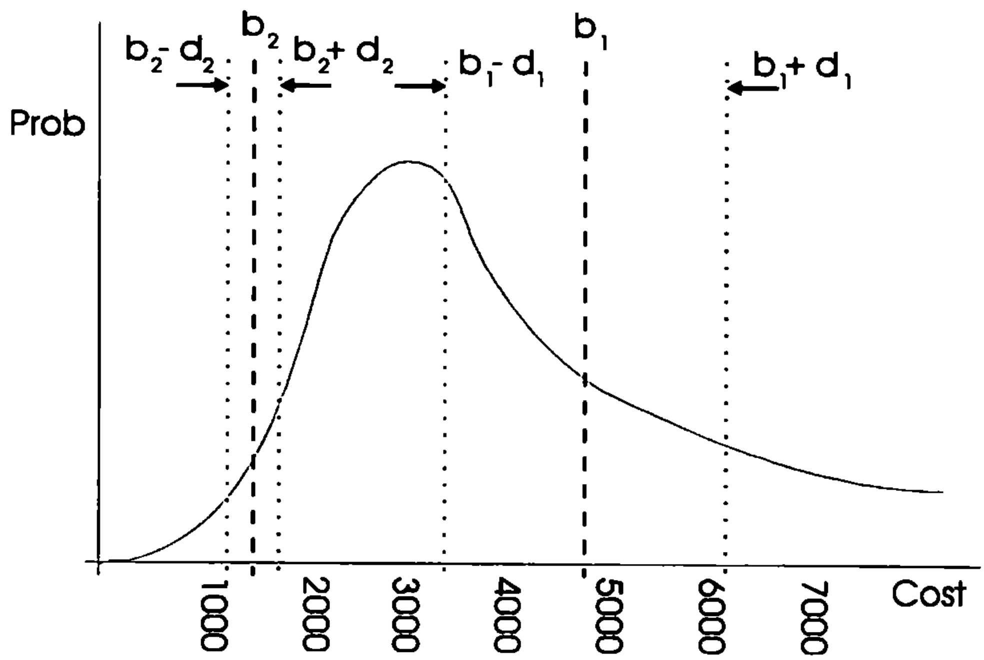


fig 6.15 A typical route cost pdf

The figure above assumes d_x = p.b_x, where p = percentage desensitivity bound factor and b_x = best cost at step x of the design process). It illustrates the difference in range set by a percentage desensitivity bound when best cost is high, b_1, and when the best cost is low, b_2. For instance, with a mean route cost of 5,000 and a broad spread of costs, if the first route cost were 10,000 and the second 1000 a desensitivity range of 30% would reject over 95% of all remaining routes. This would significantly reduce the number of routes requiring the test for the second objective.

6.7. Route Selection Influenced by Allocation Order

It is important to ensure that the desensitivity method is not dependent upon the order of the routes presented to it for comparison, since any pre-ordering process would place an unacceptable additional computation requirement on the design method. It may be argued that, irrespective of the order in which the routes are presented for evaluation, if the minimum cost route is selected and retained, then after N comparisons the overall minimum cost route will have been selected. Similarly, if each time the running best route were retained as the second best whenever a new lower cost variant were found, after N tests, the second lowest cost route would be known. This simple argument may be extended to cover the choice of routes within the desensitivity range of the best cost route.

To ensure that the order of the route cost comparisons does not cause the route selection to vary, a constraint must now be applied to the desensitivity rule. The reason for this is based on the range of the route cost, i.e. plus or minus the desensitivity factor. If the positive range is included (Best + desensitivity factor, *d*) it would be possible for a number of routes' costs to appear in succession where they were within the (Best + *d*) range and met the secondary objective. This would lead to a continual rise in the new running best cost.

While a potentially rare occurrence, this type of situation should be avoided where possible since any undue influence on the design mechanism may move the cost function in the wrong direction. The solution would be to consider only routes for the second objective where they were less than the best cost to date minus the desensitivity range. However, it was found from experimentation that the overall network designs were improved if the plus or minus range were retained. This would appear to lend strength to the argument for the use of the trade-off between cost and the second objective of hop maximisation and demonstrates the value of this technique.

If the desensitivity factor, *d*, is set to zero then this method reduces to the original minimum cost selection system. By adjusting it the optimisation process is dramatically influenced and the overall cost of networks is seen to decrease.

6.8. Searching for Optimal Routing

One of the greatest differences between the progressive, threaded route/link allocation method presented in this work and that of other researchers [Agarwal 1989], [Monma & Sheng 1986], [Mulvey 1978], [Pierre & Hoang 1990] is that in this instance links are allocated to routes as required. As far as is known, other researchers have only ever attempted to produce link based initial designs, perform routing or flow analysis on the feasible result and then perturbate the design to reduce the cost while performing rerouting if required. This places a much greater burden on the route finding algorithm since a completed feasible network has a small finite number of possible routes, limited by the topology. In this instance, where a network topology is being generated as routes are allocated, it is possible, indeed a requirement, that routes using unallocated links are costed as though those links existed. Consequently the number of possible routes is very much greater than for the fixed link scenario and the problem might conceivably be considered insolvable for useful numbers of nodes (perhaps 10 or more) unless some form of speed-up method is used.

6.9. Fast Route Search Methods

Is has been seen from the analysis of computational effort for route-finding that the larger the number of nodes to be searched through, the larger the search time, according to an exponential growth factor. However a useful reduction in the number of nodes to be searched through may be made by applying an element of foresight to the search. It would never be expected that a sensible route be made between Birmingham and London via Glasgow (for example) or from Chicago to Detroit via Los Angeles. It is clear that a bound can be set upon the reasonable paths to be taken between any two cities (nodes). Unfortunately the nature of such a bound is not entirely obvious yet any form of approximation may yield significant improvements in route search times.

Because the *Route Finder* is the core algorithm in the NDP program, it offers potentially the largest saving in computer effort. The required routes between each node pair must be recosted at each stage as the design progresses when all alternatives require evaluation. The need to re-cost is a result of two factors, firstly, as routes, and hence capacity, are allocated the cost of links may change and secondly because some links are not allowed to appear in routes. The enforced omission of links is required to ensure primary path links do not appear in backup routes. One way of achieving this is to set the link cost to infinity temporarily. The subject of penalty factors in route selection proves to be an important and powerful means for controlling route selection and is discussed later in section 12.

It is essential that any chosen route finding method should offer great computational speed, since it is the core algorithm of the design system. A number of methods are employed. Frank & Chou [1971] state 'the efficiency of any network design depends crucially on the routing procedures used in generating and operating a network'. Their 'quick' algorithm requires the determination of a list of nodes, j, connected to node i, ordered by the number of intermediate nodes. Similar shortest path algorithms are used by Gavish & Neuman [1989] and Topkis [1988].

These shortest path algorithms require that the network topology is known at the point of route selection. However the complete topology is not available during the design phase and consequently such routing methods are inappropriate. The new design method starts with a zero connected network and evaluates the cost of all possible routes. It must be possible to allocate new links to create routes and any cost associated with link setup must be included. Even though a link might not exist and require the incurring of a setup charge, there is no reason why the cost of a route between two nodes cannot be lower using some new links compared to the cost of one using only existing links. The precise nature of any route selection will be determined entirely by the instantaneous state of the topology connections, the traffic load to be routed and the cost structure in operation where the network is to be built.

6.9.1. The Elliptic Bound

In order to devise a suitably fast method of limiting the search for plausible routes through the network it is necessary to examine the data structures available in the network model and find a simple 'bounding' technique. It must be remembered that this search for routes may use links, between nodes, that are not configured in the topology at the time of the route search. Any links that might be added in order to 'create' a feasible path are allowed to be allocated if they lead to a lower cost route than was possible before their connection. Similarly, if an existing link needs to be upgraded in order that a new route be established then this is allowed if the overall cost of that new route is lower than a previous 'best'.

Since this type of method would allow the use of all possible links in a topology, for large networks (greater than 10 nodes) the number of potential new demands is proportional to the square of the number of nodes and the number of possible routes grows exponentially. It is essential that some limiting method be created to reduce the number of routes requiring evaluation to a manageable number.

A route search bounding technique has been devised based on an elliptic boundary using the route's two end nodes as the foci. Routes are only considered between nodes falling within this boundary. The advantage of using an ellipse in this way is that the calculation required to test whether any node is within the boundary requires is one addition (of the distances between the potential node and each route end-point) and a comparison of that distance with the maximum acceptable to be within the bound. The limiting distance usually being set as a simple factor of the distance between the two nodes, for example 1.5 times the node separation.

If the distance between nodes are stored in a precalculated static table in the network model, as a distance matrix for all nodes, the software implementation of this technique therefore only requires two table lookup steps, an add and a compare. This is therefore very fast in operation.

In the cost equations developed previously it was seen that it is possible to allow routes to travel further than the direct path distance if additional links with lower channel costs, i.e. of higher capacity, are used. This shows that a trade-off is possible between route distance and capacity allocation. The result of this is that the limit of a route distance using higher capacity, lower channel cost links is a fixed constant greater than the direct distance and therefore describes an elliptic region. This therefore reinforces the case for using the 'elliptic bound' route search limit . It is possible to use this as a powerful complexity reduction technique, by limiting the number of nodes searched through for plausible routes to those within an elliptic bound.

Due to the exponential growth of routing possibilities as the number of nodes increases, any reduction in the number of nodes to search through will have a significant effect on reducing the calculational complexity. This reduction may be illustrated by examining the change in the number of nodes required for route searches.

The number of route permutations with r hops from N nodes is ${}^{N}P_{r}$. Let N_e be the number of nodes within the chosen ellipse where $N_e < N$;

$$\frac{N!}{(N-r)!} > \frac{N_e!}{(N_e-r)!}$$

Let the calculational complexity for the whole problem of finding routes between all node pairs be n_{tot} using the new elliptic search bound.

$$n_{tot} = \frac{(N-1)N}{2}\frac{N_e!}{(N_e-r)!}$$

The advantage of the elliptic bounding is now apparent, since the number of nodes to search for routes between is fixed at N', so as the total number of network nodes N, increases, it can be seen that the route selection complexity remains fixed. The only overhead is the calculation of those nodes in the set N' which, as shown earlier, is small and can be performed once, with the results being stored for repeated use.

6.9.2. The Elliptic Bound in Use

Using the 'elliptic bound' technique it is possible to give the user of the Network Design Tool a set of parameters that can control both the duration of the design phase and to an extent the quality of the result. Limiting the number of nodes used reduces the route search time and hence speeds the design process. It is to be expected that the fewer nodes allowed for route searching the greater the probability that poor routes be selected. If the elliptic bound were used as a fixed factor related to the node separation the node locations would determine how many other nodes fell within each ellipse. In practice, the elliptic bound is increased from a low value until it encompasses a predetermined number of nodes. This operation is performed once before the design program runs and the results stored.

In applying the route selection procedure, all routes up to a specified Max-Hops are searched i.e. the search procedure is of the form below:

```
Route Search Procedure (A to B):
        FOR n = 1 to Max-Hops
                Find_Route with n hops between (A,B) from N nodes
```

fig 6.16 The route search with Max-Hops

Without the elliptic bound the entire set of N nodes, excluding A and B would be searched for possible routes. It would not be sensible to set the elliptic bound to be less than the number of Max-Hops since the route search would fail. The elliptic bound is therefore described as the number of nodes in addition to those required by the Max-Hops setting. For example if the Max-Hops is set to 4 and the 'elliptic bound', E_{max}, were set to 2 then the number of nodes searched for routes between would be 5. This is a result of requiring 3 nodes in addition to A and B for a 4 hop route and the further 2 nodes allowed by the elliptic bound, EB.

Bounded Route Search Procedure (A to B):

FOR n = 1 to Max-Hops

Find_Route with n hops between (A,B) from ((Max-Hops-1) + EB) nodes

fig 6.17 The bounded route search with Max-Hops

Experiments have shown (Results in chapter 7) that a point is quickly reached where increasing the elliptic search region does not significantly improve the optimisation result and yet drastically increases run time. This is to be expected given the way in which path length deviations from the linear node separation are limited by the cost savings of increased channel bandwidths, as shown earlier. The benefit of this speed-up system is relied upon for producing results in useful timescales compared to the impractical durations of exhaustive search methods.

A common application of this feature is to set E_{max} to a small value, e.g. 0, when only Max-Hops nodes will be searched, hence the design time is very short, typically a matter of a minute or two for up to 20 nodes. This is done to estimate the cost of a network very quickly thus allowing fast budgetary estimates for network proposals when only 'ball-park' figures are required and no great design time is warranted for a thorough analysis.

6.10. Spare Capacity Allocation

The presence of spare capacity in a network topology as it emerges from the design process is an inevitable result of the banded nature of capacity offerings from commercial network capacity providers, PTTs. A banded costing scheme where allocations are made in units of 1,2,4,8,16... channels is typical. In some countries, Sweden for example [Tarifica, 1993], channel speeds of only 64kbps and 2mbps are available to other European countries. The implication ~~being~~ is that if any more than 64kbps is required either multiple 64k channels must be used or a 2mbps trunk selected. There will be a breakpoint at which it is cheaper to use the 2Mbs and incur the spare capacity, typically this will occur at about 60% utilisation of the 2~~m~~Mbps. This an extreme example, but illustrates the problem of incurring spare capacity overheads due to the nature of the international capacity availability.

During the design process, as routing is established and new trunks are selected it is almost certain that spare capacity will occur at some point. The increased trunk capacity and corresponding lower unit channel costs, as discussed earlier, are sought in order to encourage the sharing of capacity and the minimisation of overall topology cost. There is a conflict between the need to increase channel sizes, reducing the unit channel costs while not incurring sufficient spare capacity to render the topology inefficient.

A trade-off in capacity allocation must be made between adding higher capacity links to benefit future route costs and minimising the allocation of capacity that will be redundant when the design is complete. The design process is therefore required to deal with the problem of to whom the spare capacity should be charged.

A simple method to recover the cost of spare capacity might be to charge each route for any spare capacity to which its allocation leads. This method poses a number of immediate problems. Since the costing is used to select routes, to charge for any spare capacity would merely discourage their allocation. The overall effect of this would be to discourage any spare capacity allocation and greatly distort the optimal route selection process. The source of this problem is that if one route were charged for any spare capacity it invoked, future routes using that spare capacity would have to make a rebate and this is not be a practical solution since the original routes were selected on the basis of their apparent cost at the time. The rebate is therefore notional.

The main alternative to charging routes for spare capacity is for the network provider to absorb the cost of spare capacity. In the context of the design process this means that spare capacity is not charged to any route, but accumulates. The implication of this is that the means by which spare capacity is 'encouraged' must be controlled. The benefit of increasing trunk capacity diminishes over the design cycle and this may be controlled using the desensitivity figure.

In the network providers favour is the fact that, as the network usage grows, or traffic patterns shift, the spare capacity may be productively utilised at a point in the future. This would be an optimistic assumption on which to rely since there is no means of ensuring that the spare capacity will be in the right place. At best it might be possible to derive an empirical figure that indicates the percentage of spare capacity productive within 3 months of installation. It is preferable to implement a method for presiding over the preference of the design system to incur spare capacity.

6.10.1. Implementing Spare Capacity Control Measures

One of the major difficulties in controlling the allocation of spare capacity is that there is no way of placing a bound on an acceptable level of overhead. This is primarily caused by an inability to judge whether spare capacity on a link may be used by the allocation of future routes or whether no more routes will be allocated to the link and the unused capacity will be wasted.

As the design process progresses from the allocation of the first to the last route there may be a change in attitude towards unused capacity. In the early stages there is a benefit to be gained by allocating extra capacity to improve future link costs. Towards the end of the design there is a potential benefit from minimising the spare capacity to reduce cost overhead. However, the issue to be resolved is 'at what point can the excess capacity be considered redundant'. Essentially, until the network design is complete and all routes have been allocated there can, by definition, be no such thing as spare capacity, which only exists because no route uses it.

The use of the desensitivity selection methods may lead to the implementation of multiple hop routes to improve future route selections and hence lead to the allocation of unused, potentially spare, capacity. In order to control the allocation of spare capacity it is necessary to consider phasing out the secondary 'maximum hop route' objective as the design process progresses towards its final route allocations.

Since the maximum hop route objective is only operational when route costs are within the cost desensitivity bound, the bound may be adjusted to control the spare capacity overhead. The way to phase out the secondary objective and to minimise capacity overhead, is to implement a system whereby the cost desensitivity is started at some value d_1 and progresses to some final value d_N for the last route to be allocated. It would be possible to force the value of d_N to be zero to minimise the final capacity overhead allocation. The starting and finishing values of d may obviously be anything the designer deems fit or has found to be beneficial for that network design.

It must also be remembered that the design phase is not complete until all the reduction methods and perturbations have been performed. For this reason, in the primary design phase the final value for the desensitivity may ideally not be zero. This is because even up to the last point in the initial design phase some spare capacity might be desirable. The best possible fit of all traffic demands into the optimal topology is unlikely to include zero spare capacity. Bearing in mind the vast number of possible network topologies the chances of co-incident topologies for both minimum cost and zero spare capacity must tend rapidly to zero as the number of nodes increases. Since the design process is not complete until after the perturbation schemes have operated the presence of spare capacity may allow the perturbation stages to reroute traffic more efficiently and reduce the overall network bandwidth.

In creating the results presented in this thesis all the work has been carried out using a linear transition from d_1 to d_N. Any other functions might be implemented but there is no apparent benefit since the desensitivity value is a bound upon the cost of routes and it was shown in section 7 how the number of routes within the desensitivity bound decreases rapidly as the route search progresses. Adjusting the value precisely over what is a sparse set of route costs is unlikely to make a significant difference. The figure below illustrates how the desensitivity value applies to a very few of the many possible route selections. The diagram shows the density of route options at each cost point, it is clear that the density drops with increased cost. Moving the desensitivity value by small amounts, up or down, does not greatly influence the individual route selections, particularly since there is a second objective to narrow down the selections even further.

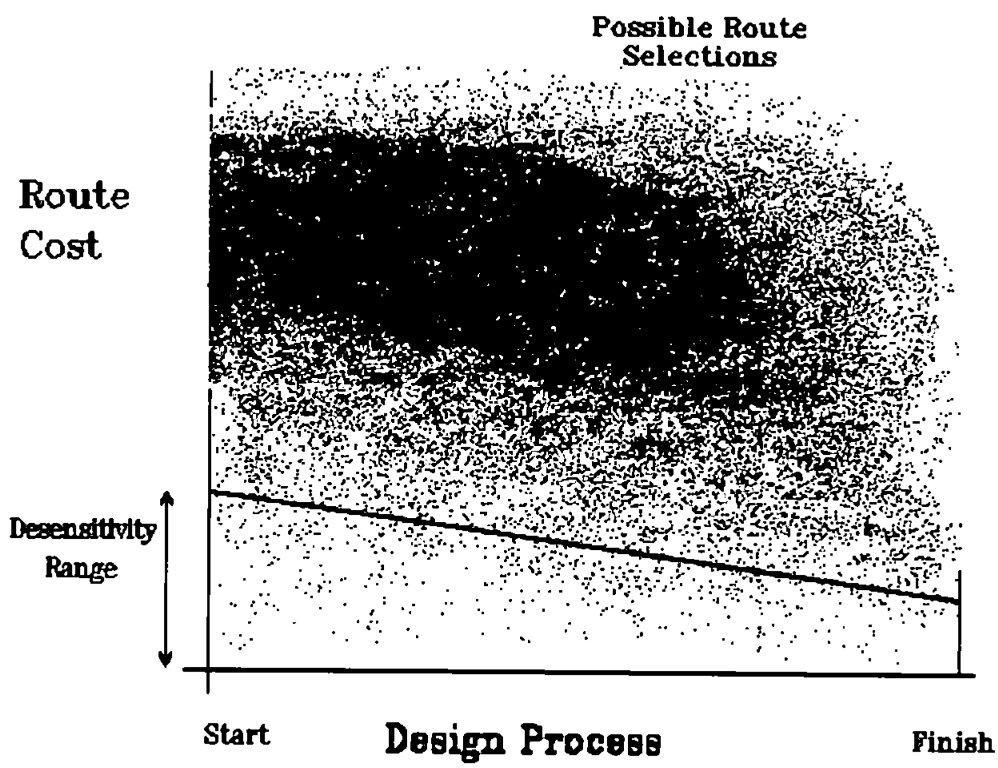


fig 6.18 The desensitivity value selects routes in a region of sparsely populated route costs

In some senses the desensitivity bound is providing a randomising effect on the route selection where a number of closely costed options exist. Experimentation with a large number of different network designs has shown that the selection of routes at each step of the design is not beneficially based on cost alone, yet the inclusion of a second objective is not the definitive solution. An element of randomisation in the second objective and repeated designs with differing desensitivity values therefore offers a viable method for attempting to 'hit' the best design.

Experimentation showed that the best network designs were achieved using either constant values or slightly decreasing values for the desensitivity. Typical settings for the start and end values of desensitivity are $d_1 = 30\%$ $d_2 = 20\%$ or $d_1 = 40\%$ $d_2 = 25\%$.

6.11. Penalty Functions

In addition to a network design method based on a topology-wide strategy it is possible to use penalty functions to dissuade specific topological features. It is possible to use many penalty functions which act in the manner of a weighted set of rules, where a cost penalty is ascribed to each topology feature that requires mitigation.

6.11.1. Link Setup Penalty

One of the major costs of a network is the allocation of individual links. There are not only fixed setup costs but administration overheads in the ordering and maintenance of each link so that it is therefore desirable to limit the number of individual links in a network. On an entirely practical point, there is a major problem in creating a network with a large number of trunks from each node, based on the provision of physical resilience to failure. In an international network, this occurs if a major node in one country has a large number of trunks emanating from it, connecting to a number of other countries. The PTT may not be able to provide separate carrier services for each trunk. There are likely to be common items of equipment in local exchanges and even common international carrier trunks. It is not unusual for international services to share cables between major national access points. Therefore a single point of failure could potentially lead to multiple network trunk failures.

It has been found from experimentation that some links allocated in the early stages of the network design process are subsequently redundant or obviously expensive in relation to the traffic carried and clearly a method of discouraging excess link allocation was required.

A further cause of the allocation of excessive links was the result of the requirement for link and node disjoint primary and backup routes. The route lockout problem, as mentioned in section 7.4 of chapter 2, can lead to problems towards the end of the design process where the primary route has taken a multiple hop path using nodes within a short distance of the direct path. The main problem stems from the backup path being required to deviate sufficiently far from the direct path to use alternative nodes that a single low capacity (sufficient to carry the demand) link is allocated between source and destination nodes. It is found that the cost savings of using higher capacities on a more diverse route are outweighed by the increased distance.

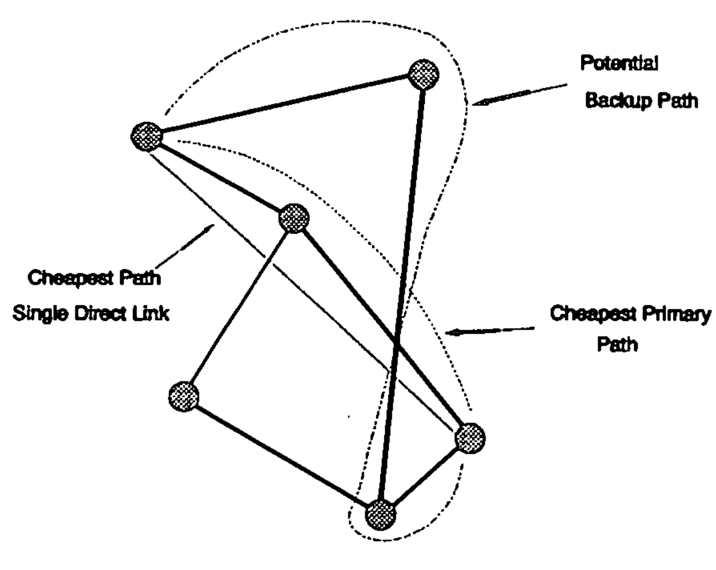


fig 6.19 The lockout of good routes may lead to unusual minimum cost routes

In order to control the allocation of new links a 'link setup' cost penalty is added to the calculated route cost. Therefore, when comparing routes, the penalty function will influence the comparison to account for the preference for using existing routes unless significantly cheaper alternatives exist.

6.11.2. Small Loop Penalty

A further penalty function investigated was one to prevent the occurrence of short loops in the primary and backup routes between two nodes. The aim being to reduce the likelihood of small loops working against the objective of forming large loops to gain cost reductions from capacity increases and per channel cost reductions. It was seen as desirable to prevent the situation arising where a low capacity path was added to provide a small loop between primary and backup.

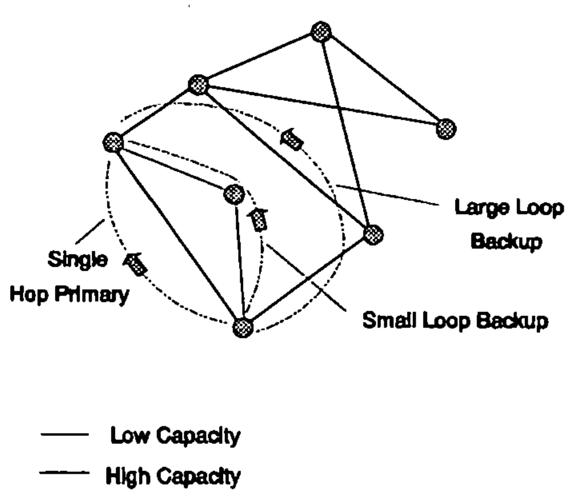


fig 6.20 The small loop penalty

The penalty function was implemented when the sum of the primary and backup hops in the routes between two nodes was less than a preset threshold. This was tried for a 3 hop threshold which would occur when the primary path was direct and the backup path had only 2 hops, thus forming a triangle.

The results showed this penalty to be of little benefit and sometimes detrimental, so its continued adoption was therefore dismissed. The reason for such a penalty being of no real benefit is found to be due to its pure reliance on the number of hops in the routes. All previous work in this project has only considered the issue of hops in routes when taken in the context of similar costs, through the use of the desensitivity bound. Therefore, while the aim of experimenting with a small loop penalty was to prevent the occurrence of small loops created by the addition of low capacity it was found that the perturbation methods implemented were more than adequate in solving this problem and this penalty was a potential waste of processing time.

6.12. The Route Finder Algorithm

Perhaps the most frequently executed software routine, of the whole design process, is the Route Finder algorithm. In this implementation it differs from the normal routing algorithms in that it makes no assumptions about the existence of links. It is normal for routing algorithms to only allow the use any links that are configured with positive real capacity values between nodes. This is because other design systems work in a manner that divorces the routing requirement from the topology selection. It is usual for a feasible topology to be created before routing takes place. This system is different in that the routing is used to build the topology.

The Route Finder algorithm employed by this design system is able to evaluate the relative costs, by the use of a figure of merit, of all possible routes between two given end nodes. The intermediate nodes allowable for use in a route are those that fall within the elliptic bound of the two end nodes. A figure of merit for a route is used to evaluate the relative preferences for each possible path. It is made up of the cost of the route plus any penalty figures that must be added. The cost is based on the sum of all pro-rata costs for each channel of each link in the route.

$$f_{ij}^{n} = \sum_{l=1}^{L} (D_{ij}.c_x + s_x)$$

where f^n_{ij} = figure of merit for n^{th} route between node i and node j;

L = number of links in route from node i to node j;

c_l = cost per channel of the l^{th} link;

D_{ij} = traffic demand, in channels, between node i and node j;

s_l = setup penalty for each link; $s_l = 0$ if link l established.

If a link is required but not present between two nodes then it is initiated at the smallest available size to just carry the traffic demand being routed. If a link is required that is configured with insufficient capacity then it is upgraded to the next available capacity to just carry the existing traffic plus the new load.

It must be stressed that the addition of capacity for each route evaluation is made only temporarily for the duration of the individual route's evaluation. It is deleted from the capacity allocation matrix after the route under test has been evaluated, in readiness for the next test.

A typical implementation is of the form:

```
FOR all possible routes between i and j
        begin
        Using only those nodes within the ellipse
        Form a potential path P, of capacity D
                begin
                Allocate capacity D to each link in path P
                Calculate the figure of merit for path P
                Deallocate capacity D from each link in path P
                end
        end
```

fig 6.21 The Route Finder algorithm

6.12.1. Node Lockout

When searching for primary routes between two nodes all possible intermediate nodes are examined and the minimum cost, or minimum figure of merit, path is selected. However, when searching for a backup path it is vital that no nodes used in the primary route are included in the backup. For this reason it is imperative that a means of 'locking out' nodes is available. A list is maintained of any 'locked out' nodes and is used to exclude any invalid nodes from the elliptic bound calculation when nodes are selected for the route-finder algorithm.

6.13. Network Cost Reduction

It is accepted that, though using an improved algorithm, the desensitivity based route selection is based on a partially greedy objective, the initial topology may therefore be sub-optimal. In order to minimise the network cost where possible, one or a series of design perturbation system may be employed.

The typical cost progression through the first to nth stages of the optimisation will therefore be typified by a phase of cost rise due to link allocation followed by a cost reduction phase. The potential for cost reduction will depend upon the particular reduction methods applied though there is no guarantee that the lowest final topology cost will result from the lowest cost initial design.

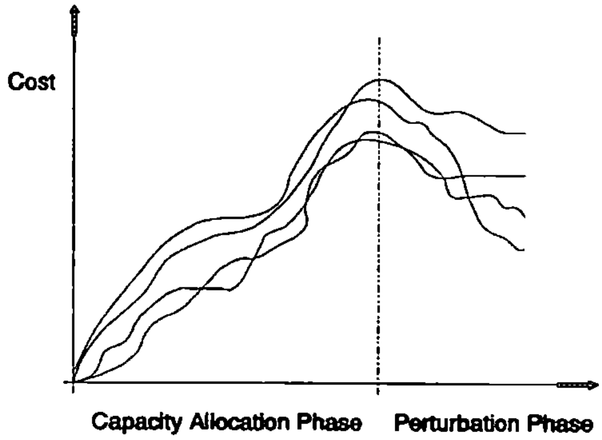


fig 6.22 Topology cost function over the design cycle

Figure 6.22 shows the typical cost function through the design and perturbation phase. It is important to notice that the cost of the topology at the end of the allocation phase does not directly relate to the cost of the topology following perturbation. This becomes evident after viewing the costs of the results from a number of design processes.

6.13.1. Cost Reduction methods

The cost reduction methods employed are based on a simple core algorithm:

i. select a trunk for capacity reduction;

ii. reroute any impacted paths;

iii. if cost improved then keep modification, else restore original.

In developing the new design procedures it was found that an improved execution speed of the perturbation phase would result from the storage of the route data for each node pair. This allows any analysis method to interrogate the model and determine which particular routes are using each link. This then offers the potential to reroute individual traffic demands as deemed necessary to achieve a 'best fit' routing within the available links.

All the perturbation schemes implemented rely on the model retaining full route information. Without the large amounts of data held to describe each route the new route based reduction methods would not be practical. Furthermore, given the highly theoretical nature of many papers on this subject [Frank & Chou, 1971], [Gerla & Kleinrock, 1977], [Esfahanian & Hakim, 1985], [Balakrishnan & Graves, 1989], [Gavish & Neuman, 1989] it is perhaps understandable that routes have not before been regarded on an individual basis since both the notation and the storage requirements become cumbersome for even small numbers of nodes. For example, 10 nodes require a 10x10 matrix to describe each possible connection and a maximum of 10 hops per route would increase the requirement to 10x10x10. Only as larger computer memory systems become commonplace will this system of holding route data be applicable. The IBM-PC used for this research was limited to approximately 256kbytes of variable storage after the allocation of program and graphics memory. Larger personal workstations such as those running the UNIX operating system and offering large amounts of virtual memory would be required to make this method applicable to problems with more than approximately 16 nodes.

6.13.2. Spare Capacity Removal

The title of this reduction technique refers to the method by which capacity adjustments are sought. The capacity of a link is analyzed for potential adjustment if a single route can be found to have caused the capacity to be upgraded from one value to the next higher e.g. if a 128kbps trunk has 128kbps allocated and a 64kbps channel were added the next available trunk speed of 256kbps would be required, leaving 64kbps of capacity unused. The aim of this reduction method is to minimise the difference between that capacity allocated and that assigned to specific routes.

If the capacity of any link may be dropped to a lower value by the removal of a single route and the overall spare capacity of the network reduced then a cost saving is possible. There will only be a lowering in overall network cost if the saving of the spare capacity on the links over which the deleted route is not outweighed by any capacity used in rerouting the demand and any further spare capacity entailed as a consequence.

```
FOR all trunks
        begin
        trunk_routes = COUNT number of routes on trunk (i,j)
        FOR all trunk_routes (i,j)
                begin
                Remove route (x,y)
                Lockout route (x,y)
                Search for new route (x,y)
                IF cost (new route (x,y) < old route (x,y) then
                        Allocate new route (x,y)
                ELSE
                        Allocate old route (x,y)
                end
        end
```

fig 6.23 The spare capacity reduction technique

6.13.3. Single Route on Link Redundancy

A simple perturbation scheme that was found to be highly effective under certain circumstances is based on examining the number of routes per link. If any link has just one route using it then there is a high chance that the route in question may be deleted and successfully reassigned to another path. The primary reason for this test being useful is that it can remove links that are configured early on in the design phase but never augmented as the result of 'better' links becoming available for other routes at a later stage. When the initial design phase is complete these single path links may often be deleted and their route profitably reassigned to an alternative path. This simple 'single route' redundancy test is therefore a useful quick method. A benefit of this test is that it is very simple to implement since it only has to test each link once and the disturbed path is rerouted using a single iteration of the Route Finder algorithm from the main design procedure.

The test for such an alterative route is performed in much the same way as the spare capacity reduction system, i.e. it is necessary to first delete the route under test, inform the Route Finder that it is to be ignored and then search for all possible alternate routes and select the one with the lowest figure of merit.

```
FOR all trunks
        begin
        trunk_routes = COUNT number of routes on trunk (i,j)
        IF trunk_routes = 1
                begin
                Remove route (x,y)
                Lockout route (x,y)
                Search for new route (x,y)
                IF cost (new route (x,y) < old route (x,y) then
                        Allocate new route (x,y)
                ELSE
                        Allocate old route (x,y)
                end
        end
```

fig 6.24 The single 'route on link' reduction technique

6.14. Software Development

The Network Design Tool developed during the completion of this project was written in the Turbo Pascal language on an IBM-PC. The software was split up into logical sections, where the design algorithms were separated from the graphics, management, file-handling and basic analysis primitives such as Route Finder and the costing functions.

The software development programme consisted of two cycles each made up of three phases. The first cycle involved the creation of a simple graphics user environment that was used to display nodes and trunks and execute simple design algorithms. The three phases of the first cycle were:

i. graphics environment development;

ii. development of network analysis software primitives;

iii. development of network design algorithms.

In the first development cycle the software was written largely as a proving ground for ideas to ensure that basic design processes would operate at acceptable speeds on a personal computer. This first cycle was also important in that it allowed for the development of the various software techniques required to model computer networks and design requirements.

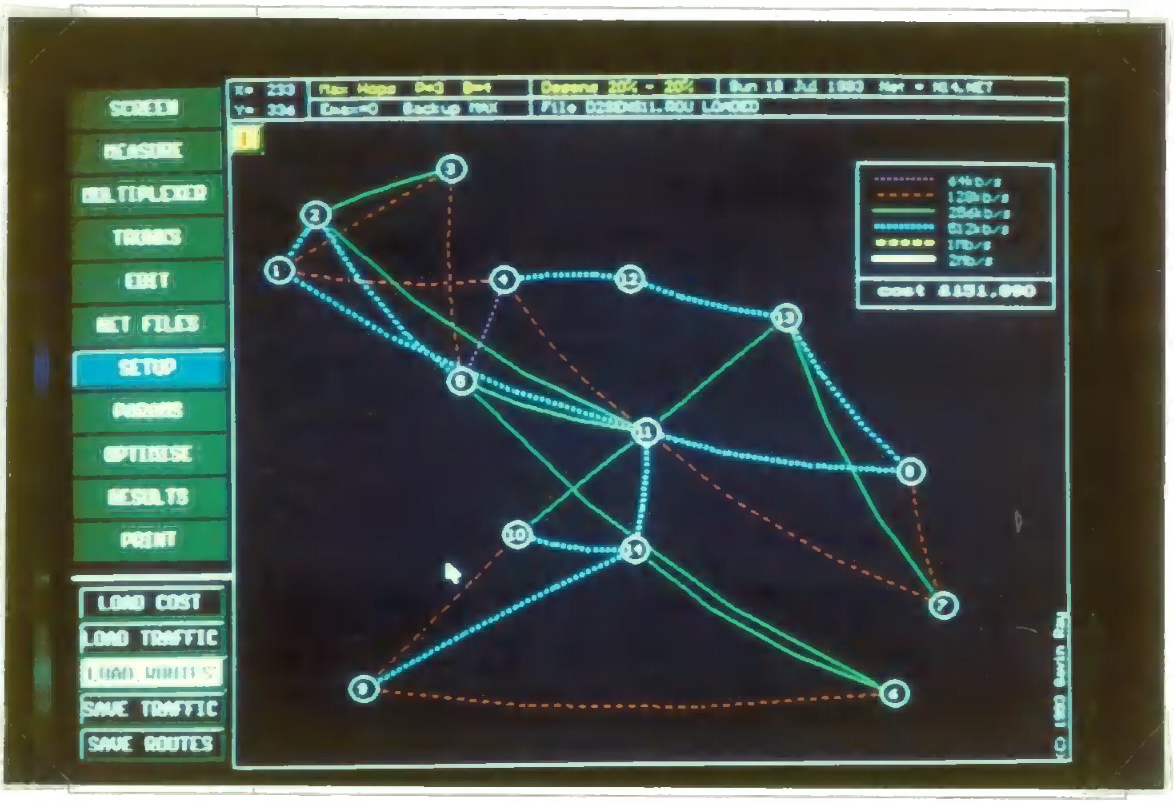


fig 6.25 The Network Design Tool graphics window

The figure above shows the format of the Network Design Tool user graphics window. The mouse pointer may be moved to select the various setup and design options from a series of options on the left. The main window shows the network topology under design. The top bar contains status information, primarily as time and date of analysis, current operation in progress, topology name and parameter settings such as maximum hops allowed for routes and ellipse size. The key shows the capacity indicated by each coloured link between nodes.

The second development cycle involved the complete rewriting of the software to incorporate new methods.

Once the first software development cycle was complete it was clear that a personal computer would offer sufficient processing power to allow the design of networks with up to 16 nodes using link based design algorithms within a computational period of a few minutes. This matched the requirements for the design of European networks with up to 16 nodes and the development continued into a second cycle. The software was then rewritten over a period of about one year to increase the sophistication of the design algorithms, analytical functions and graphical interface. The new system implemented the full route based design algorithms and the high speed Route Finder analysis methods using the elliptic bound.

6.14.1. Description of the Network Design Tool, the 'Optimisation Engine'

The Network Design Tool created for this project was named the 'Optimisation Engine' in order that it could be uniquely referred to externally to the research project. The software structure is shown in the figure below.

Graphical Interface

- Window (open,close,move)
- Draw Menu
- Select Menu Item

File Management

- Read File (link, node, route, cost, demand)
- Save File (link, node, route, cost, demand)
- Edit File (link, node, route, cost, demand)
- Create Report File
  - List Link Usage
  - List All Routes
  - List All Costs

Parameter Control

- Set Max Hops per Route
- Set Ellipse Range
- Set Link Penalty Factors

Analysis Primitives

- Find Route (from A to B)
- Count (Routes on link A to B)
- Count (Channels spare on link A to B)
- Cost (Route from A to B)
- Cost (Link from A to B)

Design Algorithms

- Perform Link Based Design
- Perform Route Based Design
- Perform Redundancy Removal
- Perform Spare Capacity Reduction

fig 6.26 The 'Optimisation Engine' structure the number of nodes in the overall network. Hence the design times do not extend as dramatically with increasing network size as they would were all nodes tested in each route.

The graphs below represent the run time data shown in tables 7.1., 7.2 and 7.3. They show the results for a number of design exercises that were carried out for differing size networks with required maximum path lengths of two and three hops.

In the following tables the abbreviations are as follows:

E_{max} = elliptic bound limit;

(maximum number of nodes for route search in addition to hop limit);

P = maximum number of hops in the primary path;

B = maximum number of hops in the backup path.

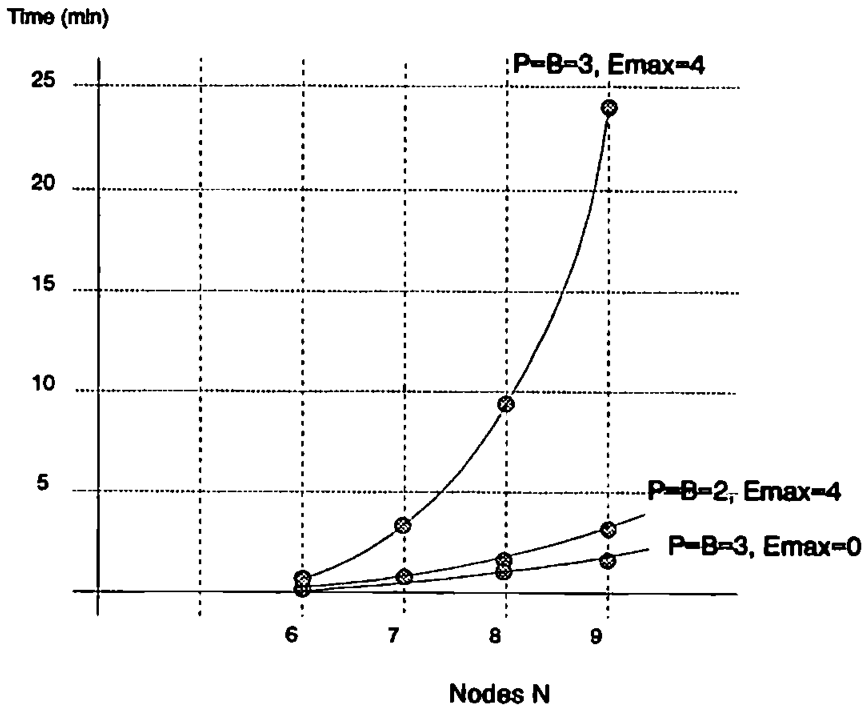


fig 7.2 Graph to show the run time for different sized topologies

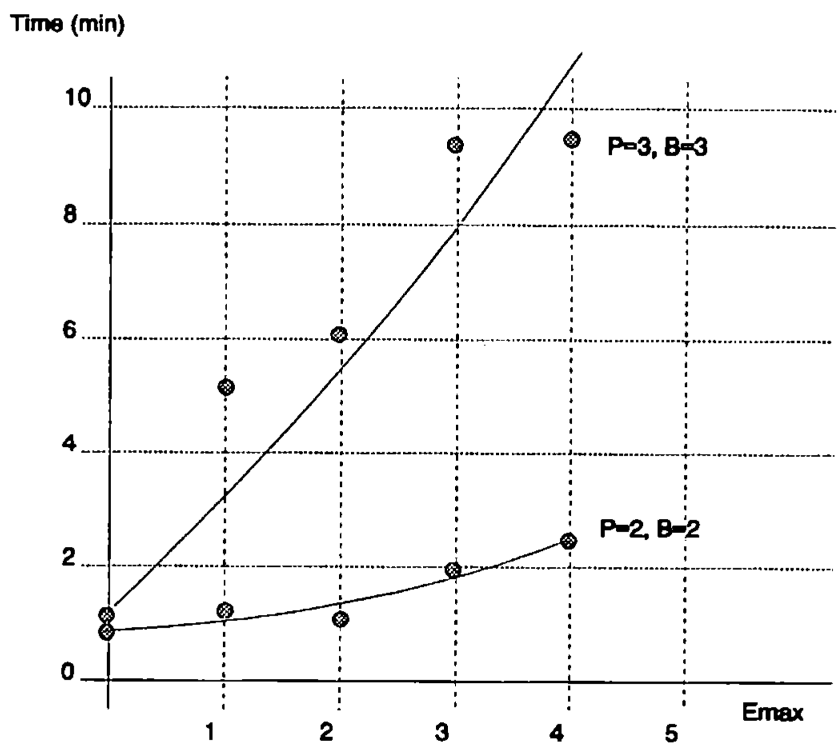


fig 7.3 Graph to show the design times for 8 node topology

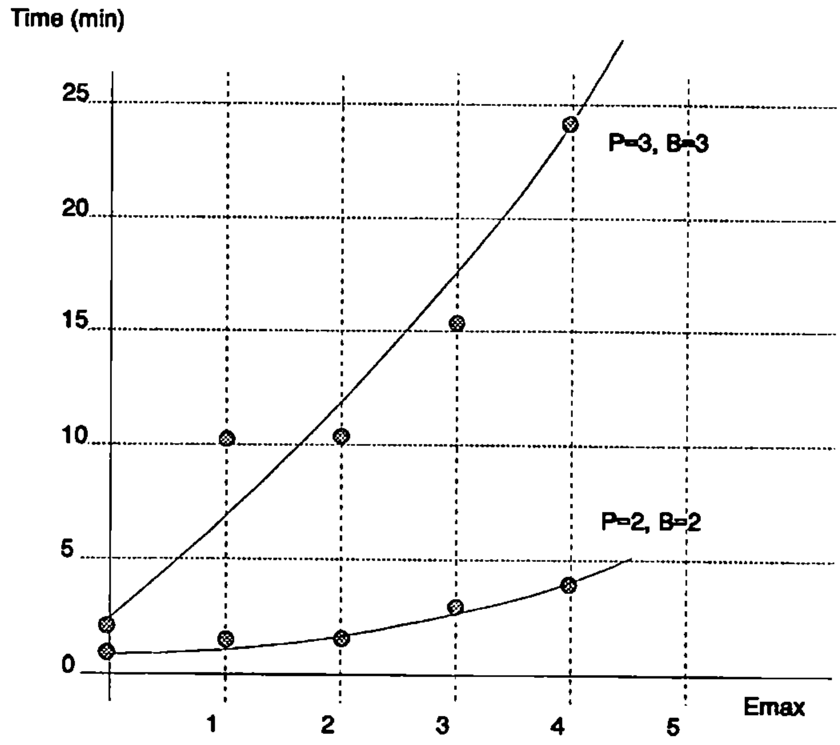


fig 7.4 Graph to show the design times for 9 node topology

The routes are searched for in order, starting with 1 hop paths increasing up to P_{max} maximum primary hops and B_{max} maximum backup hops. The elliptic bound is described by the value E_{max} which represents the maximum number of nodes between which to search for routes, in addition to those allowed for by the P_{max} and B_{max} parameters. The number of nodes used in the route search is therefore a maximum of $P_{max} + E_{max}$:

e.g. if $P_{max} = 3$ and $E_{max} = 2$, then the elliptic bound is extended to encompass the 5 nearest nodes to the end points falling within an ellipse.

In terms of computational complexity this therefore implies that the work required to search for maximum 3 hops routes with $E_{max} = 2$ is equivalent to searching for 5 hop routes with $E_{max} = 0$.

In the following tables 7.1, 7.2, 7.3 and 7.4 the design times and network cost are shown for a range of design tests. Using the following topology parameters:

i. a number of randomly placed nodes;

ii. a cost matrix approximately based on UK type setup and distance related charges;

iii. a traffic matrix of small, randomly selected internodal demands.

A number of design exercises were performed with the number of nodes, N, ranging from 6 to 9 to illustrate the increase in design time with increasing node count. The hops indicated show the primary and backup path length limit, which was kept equal in each test, though they may be set to different values.

The tables also show the effect of extending the elliptic bound, E_{max} from 0 to 4. The design times may be seen to extend with increasing E_{max}.

Nodes	Hops	E_{max}	Cost	Design Time Secs
6	2	0	59,210	15
		1	54,246	28
		2	54,246	26
		3	54246	26
		4	54,246	26
	3	0	54,246	27
		1	54,246	72
		2	54,246	72
		3	54,246	72
		4	54,246	72

Table 7.1 N = 6. Cost and Design Time

Nodes	Hops	E_{max}	Cost	Design Time Secs
7	2	0	90,426	24
		1	77,387	55
		2	77,387	51
		3	77,387	61
		4	77,387	61
	3	0	77,387	54
		1	77,387	163
		2	86,647	174
		3	86,647	174
		4	86,647	175

Table 7.2 N = 7, Cost and Design Time

Nodes	Hops	E_{max}	Cost	Design Time Secs
8	2	0	116,114	41
		1	100,219	94
		2	108,114	84
		3	108,114	109
		4	108,114	132
	3	0	100,219	93
		1	102,577	339
		2	101,382	370
		3	91,212	565
		4	91,212	565

Table 7.3 N = 8, Cost and Design Time

Nodes	Hops	E_{max}	Cost	Design Time Secs
9	2	0	139,834	63
		1	132,294	146
		2	134,663	131
		3	134,663	184
		4	134,663	233
	3	0	132,294	146
		1	129,570	612
		2	127,388	640
		3	116,923	1012
		4	116,923	1437

Table 7.4 N = 9, Cost and Design Time

The results in tables 7.1, 7.2, 7.3 and 7.4 show that the optimal routes are not found when E_{max} is 0, but more generally when E_{max} is set to 2, 3 or 4. As the number of nodes, N, increases from 6 to 9 the optimal value of E_{max} takes on various values, but rising from 1 to generally 2 or 3.

In examining the design times for each parameter it is noted that while the results for E_{max} = 3 and 4 are generally of the same cost, it is the lower setting that yields the speed advantage. The rapidly rising slope of the time function indicates that there is considerable time advantage in using the lowest possible setting of E_{max}. The tradeoff is a risk of a design of poorer quality.

Since this network design method is required to analyse networks of greater than 7, 8 or 9 nodes with correspondingly large design times it is first necessary to identify the optimal values for E_{max} so that analysis of the other parameters may be conducted in ~~un~~reasonable periods of time.

Due to the large number of other parameters it is possible to adjust in the design process to limit the variability of results, these tests were conducted with fixed and equal values for maximum primary route hops, P_{max}, maximum backup route hops, B_{max} and desensitivity.

It can be seen that using an elliptic bound, E_{max} with values of 2 or 3 can significantly reduce design times while still leading to good design results. For example, in the 9 node network, with a 3 hop route search, when $E_{max} = 4$ the design time was 1437 seconds yet reducing E_{max} to 1 more than halved this to 612 seconds. Comparing this with the 8 node network, the same reduction in E_{max} lead to a reduction in design time from 565 seconds to 339 seconds, again a significant reduction, but not as great relatively. As the number of nodes increases, the effect of the E_{max} route search reduction becomes more significant.

A value of Emax = 1 gives very rapid designs, within less than 10%, typically, of the lowest cost topology found using an extended bound. This can therefore be considered useful for determining approximate topology cost where speed is critical. It is likely to be particularly useful during the initial problem specification where no idea of cost is known, a rough cost guide can be quickly provided and influence the extent to which traffic requirements might be expanded or contracted based on any given budget constraints. This is typical of the requirement from the initial sales interface between the network provider and the customer who has an outline requirement which will be finalised once typical costs are known.

7.4. Testing the Desensitivity Factor

A network of 14 randomly placed nodes was used for the testing of the desensitivity factor. The node placement ensures a range of internodal distances with no symmetry about any axis. This is intended to mimic typical physical node layouts that are equally unlikely to be regular in any way. The choice of 14 nodes was based on the need to examine realistic sized problems without extending the analysis time unduly.

Of equal importance to the node locations is the internodal link cost structure and the traffic requirement. The link costings were made approximately proportional to the UK PTT costs. A typical link setup charge and distance related tariff was applied for each of the 6 available trunk capacities, 64kbps, 128kbps, 256kbps, 512kbps, 1Mbps and 2Mbps. When 'fractional bandwidths' are available (i.e. multiples of 64k) it is common to find a typical capacity doubling progression where a 100% capacity increase has a 65% cost premium.

The traffic requirement for the 14 node network was also structured to offer a random distribution of channel requirements between both local and distant nodes.

All final testing was conducted on a 80286 processor based PC with approximately half the processing power (processor clock speed x instructions per cycle) of the 80386 used for development.

In order to assess the effectiveness of the desensitivity factor the 14 node test network and associated load and traffic files were the subject of a number of design exercises. The aim was to determine whether the expected trade-off between link capacity increases and extended path length were justified. For the cost desensitivity factor to be effective the cost structure needs to be such that increasing the channel capacity lowers the cost per channel.

In addition to the cost structure applied to each link there is a penalty cost function that may be used to discourage the allocation of an excessive number of links. This provides a means for including in the design process the real world cost of trunk implementation such as trunk administration, management, maintenance and interface costs. The test results were repeated to analyse the effects of the desensitivity factor and the link setup penalty in tandem.

7.5. Link Setup Penalty Function

The table 7.5 below shows how the network cost varies following two different design scenarios, the first cost column shows the result~~ant~~ing cost of the network using no link setup penalty. The second cost column shows the result of using a setup cost penalty of 2000 pounds plus a further 20 pounds per kilometre of trunk. All costs can be considered to be in pounds, though the particular currency is irrelevant, only the cost relationships between the different topologies.

P	B	E	Dmin	Dmax	Cost (Setup = 0 km = 0)	Cost (Setup = 1000 km = 20)	improve-ment
3	3	2	0	0	161,191	146,291	9.2%
3	3	2	10	10	161,385	145,873	9.6%
3	3	2	20	20	158,177	141,002	10.9%
3	3	2	30	30	158,177	139,284	11.9%
3	3	2	40	40	158,796	141,711	10.8%

Table 7.5 Topology Cost with and without the Link Setup Penalty

The results clearly show how the penalty function benefits the overall network cost. In common with other results it is seen that the overall cost improvement is also related to the cost desensitivity bound. As the desensitivity value rises from zero to 30% the overall cost reduces. At a desensitivity figure above 30% the cost then starts to rise again. This is approximately the effect that was predicted in chapter 6 where it was recognised that the route selection procedure would tend to increase the allocated capacity, using fewer trunks. Up to a limit this is beneficial since the increased capacity has a lower cost per channel and contributes to a reduction in overall network cost. The rate of capacity increase is partially determined by the desensitivity factor but it must also be acknowledged that there is a very complex relationship between this and the traffic load, trunk cost structure and the node layout.

From the results above it may be estimated that cost desensitivity settings in the range 20% to 30% appear to be the most useful initial values for starting any new analysis. Subsequent adjustment may then be made and the resultant topology costs compared to the expected 'good' starting topology costs.

7.6. Visual Appearance of Topology Designs

The following results represent a number of topologies from the design scenarios above. With reference to the discussion of visual appeal of topology in chapter 6 there are a number of features of the design results that become apparent.

It may be surprising that a large number of designs have been produced within a narrow cost band, some 10% between worst and best, and also having very different appearance. However, there are some common features to the designs. The first is the difference between peripheral trunks and central trunks, as indicated in 8.2. A common theme of rim and spoke trunks is clear, that is there appears to be a number of trunks linked in a chain around the edge of the network, usually of a lower capacity than the core trunks. This is typical of the travelling salesman tour result, where large arcs are produced between nodes. This offers the most economic solution where routes may 'bunch' together and benefit from the lower channel costs of common paths and high capacity links. The feature of the rim is that is offers the minimum number of trunks to connect together the peripheral nodes and hence reduces the costs proportional to the number of links, i.e. the setup cost.

Another surprising result is that two networks can occur with costs differing by only 0.3% yet having very different structure. Referring to the examples in appendix A, taking as an example network designs A2 and A6:

Cost A2 = £145,873

Cost A6 = £145,417

Difference = £456 = 0.3%

The absolute cost difference is not certain to be the dominant factor in the selection decision between topologies A2 and A6, the scale of the costs will determine whether they are critical. If these costs are the monthly circuit rental for a small network then the cost decision would give way to other preferences between the topologies, such as selecting the topology with the lowest trunk count to reduce management and maintenance overheads. However, if the costs shown are the daily circuit rental for an international network, then a cost saving of the order of £180,000 per year is indicated. Since this sum equates to the employment of possibly four or five staff the decision is likely to be cost critical.

The upper part of each network, A2 and A6, above node 11, is identical in both trunk and capacity detail, yet the lower sections differ greatly between the two. There are 7 trunks that differ in both connectivity and capacity. This perhaps best illustrates one of the greatest problems in network design, the vast number of variables and very closeness in cost of many of the possible topology costs makes differentiating between the good and bad attributes extremely difficult. In trying to evaluate the 'good' and 'bad' topologies there is a requirement for a method by which to formulate qualitative judgements.

The empirical test results shown in appendix A demonstrate the existence of many topologies with closely matched costs. To guide the selection between a number of similarly priced solutions a consideration of the effects of shifts in traffic profile is useful. Since traffic shifts are likely as some network customers either change premises, increase loads or even cease custom, the topology that can show the greatest insensitivity to load movement is to be preferred.

An even distribution of spare capacity is one possible basis for measuring load insensitivity. However, there is a further point of which to take account. Consistent with the more common spoked hub type of network, there is an increase in traffic levels towards the centre (of demand) in networks. This leads to a capacity contour from the edge inwards and a corresponding spare requirement contour following a similar pattern. The larger the trunk capacity the larger the spare capacity desirable to cope with growth and traffic shifts. Hence it is most common to express spare capacity requirements in percentage terms, relative to trunk capacity, rather than in terms of absolute channels. It is not possible, or at best highly unlikely that a completely even spare capacity distribution will occur, yet the lowest variance of spare capacity across a topology is a suitable objective.

Each topology created by the network design tool is shown on the screen in graphical form, together with a textual report to show the numerical data relevant to the design. The report details all routes on each link, all links used by each route and the spare capacity on each trunk. A single utilisation figure is also given to indicate the ratio of capacity allocated versus the capacity spare. It is thus possible to judge the merit of a number of designs by the detailed analysis of the spare capacity on each trunk or coarsely judge them on the basis of the spare/allocated capacity ratio.

7.7. Comparative Analysis of 14 Node Topologies

The simple aim of a network design exercise is likely to be the search for the minimum cost topology. However, since cost is just one of the metrics applied to network designs, the minimum cost topology might be the first to be examined but is not certain to be the design selected. The table shows a range of other features by which the topologies may be judged. It has been stated in previous chapters that the minimum cost network only satisfies the single set of design inputs, in particular the traffic demand.

In the table below the results of a range of 14 node design tests are shown, the network topologies are shown in Appendix A. The table header abbreviations are as follows:

Topology = name assigned to the test topology file;

Desens = desensitivity;

Util = utilisation;

P/B = No. of Primary channels allocated / No. of Backup channels allocated.

Topology	Desens (%)	Cost (£)	Spare Value (£)	Cost Util (%)	Capacity Util (%)	P/B
D2SENS-0	0	151,476	30,969	79.6	76.6	1.57
D2SENS-1	10	145,417	33,834	76.7	72.6	1.11
D2SENS-2	20	151,083	22,270	85.3	83.0	1.11
D2SENS-3	30	140,556	22,153	84.2	81.1	0.90
D2SENS-4	40	144,553	18,166	87.4	83.9	0.74
JUN29-1	0	146,291	21,116	85.6	83.8	1.38
JUN29-2	10	145,873	30,347	79.2	75.2	1.31
JUN29-3	20	141,002	23,848	83.1	80.0	1.37
JUN29-4	30	139,284	20,480	85.3	80.2	1.20
JUN29-5	40	141,711	22,428	84.2	81.0	1.08

Table 7.6 A range of 14 node network solutions

Any real network is expanding over a period of time as user requirements naturally tend to increase and the selection of a suitable topology must account for this. It is feasible that the lowest cost topology has a very high capacity utilisation and little room for traffic growth, whereas, for a small increase in cost an alternative design may offer a substantial increase in growth potential.

In chapter 5 is was indicated that the backup capacity requirements are approximated since it is not possible to make a precise judgement on the backup channel requirement while the topology is incomplete. In order to judge how much capacity has been allocated the design software produces figures for primary and backup channels, together with the mean number of hops for primary and backup paths.

The table below shows the backup capacity ratio for each topology.

Topology	Desens (%)	Cost (£)	Spare Value (£)	Backup Capacity Ratio
D2SENS-1	10	145,417	33,834	0.60
D2SENS-2	20	151,083	22,270	0.67
D2SENS-3	30	140,556	22,153	0.66
D2SENS-4	40	144,553	18,166	0.59
JUN29-1	0	146,291	21,116	0.57
JUN29-2	10	145,873	30,347	0.62
JUN29-3	20	141,002	23,848	0.62
JUN29-4	30	139,284	20,480	0.58
JUN29-5	40	141,711	22,428	0.60

Table 7.7 The backup ratios for various 14 node networks

The results show that the observed backup capacity ratios have a low variability, 10% between minimum and maximum values. This therefore supports the use of an approximation method in this instance since there is no wildly fluctuating allocation to disrupt the quality of results. Since the backup capacity varies from one topology to another it is necessary to judge its merit in each case. The two extreme cases of backup capacity allocation are either an excessive allocation causing over-expenditure or an insufficient allocation leading to failure to carry all circuits under the worst case single failure condition.

The backup expenditure issue is significant in that the over-allocation of backup capacity could nullify any cost savings otherwise made by a 'good' design and the under-expenditure problem would threaten the commercial viability of the network. In judging the merits of a number of topologies it is therefore possible to either select the topology that maximises backup capacity in favour of failure resilience or to minimise overall expenditure. The backup capacity ratio is therefore most useful in selecting between two closely costed topologies, where the cost difference is negligible, it is possible to select the topology on the basis of the maximum backup capacity allocation. This therefore prefers a topology with a greater proportion of capacity allocated to backup traffic provision and hence a 'safer' topology.

CHAPTER 8

Conclusion

8.0. Introduction

The network design problem is complex, yet primarily the problem may be broken down into two logical sections. There is a clear distinction between an access network and a core transportation network. The access network is considered to be the portion that collects the multitude of data inputs and presents a unified data stream at a single point to the core network. The core network is therefore described as that part of the network which transports the submitted data across substantial distances, as required between the many access points.

The distinction between access and core functions is useful for two reasons. Firstly it separates the complexity of the access concentrator location problem from that of the concentrator inter-connection problem. Secondly there is a clear physical differentiation between the two problems in the commercial environment and it is therefore unnecessary to tackle the problems simultaneously.

8.1. Conclusion

The access design problem is largely created by the difference in traffic statistics due to the different type of communication session conducted by the various classes of network user. The optimal mix of users is sought to create a balance of light and heavy traffic demands and lead to acceptable performance figures, which are seen in terms of character delay. The visual representation of the expected character delays, that result from various user combinations on access concentrators, has resulted from the development of a new means for estimating the likelihood of given delay figures. This is valuable in the budgeting for equipment and assisting the engineering of the access network in order to meet service level agreements between customer and network provider. It is also of great assistance in estimating the performance impact of changes in traffic due to rises in the number of customers and increases in traffic volumes.

The inter-connection of concentrators with a high capacity core network on a national scale requires considerable investment and offers the opportunity for great cost savings by careful design. The core network design problem was tackled separately and an effective TDM core network design and optimisation system has been developed. Good network designs have been generated, to a quality suitable for use by a major international network provider. The new method offers an alternative solution to a problem previously only tackled using systems based upon linear programming techniques and requiring considerable mainframe processing power. Two new techniques allow similar sized problems to be tackled on medium power personal computers. The first new technique, called 'cost desensitivity', improves the quality of designs using greedy algorithms by sacrificing short term gains for a global topology objective. The second new system used in route selection is called 'elliptic bounding'. This greatly reduces the execution time of searches for feasible paths by forming a logical limit to the number of nodes between which routes are tested. Using the simple equation of the ellipse the test itself adds an insignificant overhead to the overall calculation. The speed improvements of the elliptic bound are related to the size of the bound, the greater the speed the less likely are the routes to be optimal. However, network designs can be constructed rapidly using fast search methods and cost estimates given to within typically 10% of a complete solution's final cost using a slower, more comprehensive search.

The designs can be shown to satisfy all given constraints and design costs have been exhibited within limits acceptable to the end user. The results are generated on a medium power IBM-PC within timescales (typically 10 minutes to 5 hours) acceptable to the designer who is given full control over network parameters and the opportunity for manual adjustment. Cost optimality is not guaranteed but, in order that differing network designs may be compared qualitatively, detailed figures are given showing the breakdown of both financial and capacity utilisation for all network loads.

Perturbation methods have been developed successfully from basic techniques outlined by other researchers in the field of packet switched network optimisation . The new methods have been shown to provide useful cost savings when applied to real TDM network designs.

The disadvantage of this network design system is that the approximation method used for backup capacity allocation has no guaranteed efficiency. Since it is not a definitive capacity allocation system there is a need for a further design stage to be implemented. The backup capacity can only be determined from a completed topology design where the effect of failing each trunk in turn can be analyzed. From this it is then possible to calculate the resultant additional capacity required to handle all rerouted traffic. This requires the future development of a separate design routine for backup capacity verification.

The consistency of the design quality gives an indication that the design methods are soundly based. The new network design system promises to be invaluable, allowing network providers to supply quick tenders for new contracts, by greatly reducing the time taken to develop viable cost estimates. A further important application of such a rapid design system is in long term financial planning. International networks cost many millions of pounds and there are, by necessity, considerable efforts required in advance of such expenditure for ensuring a valid commercial basis for each network component. Long term planning is based on market strategy and sales forecasts and these are used to create traffic volumes for which topology costs are required. The planning process is based on a series of iterations between marketing, sales, finance and engineering departments and does not require topology designs of ultimate minimum cost; within 5% or 10% close of optimum is entirely acceptable given the lack of precision in the traffic forecasts.

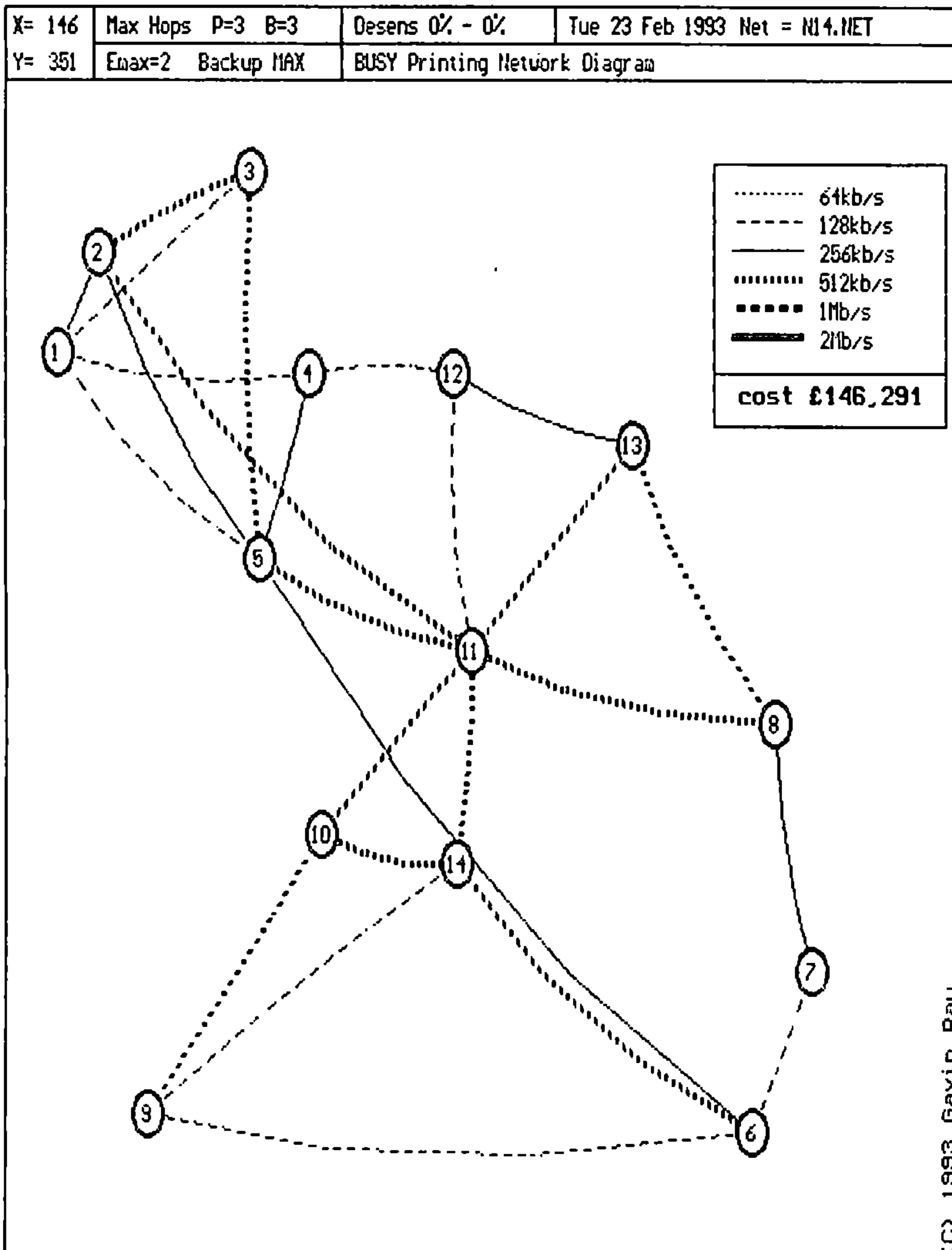
An opportunity for further development exists in the area of perturbation schemes. It has been clear from the review of a number of topologies that the design engineer can sometimes, though not consistently, see a design improvements that the current software is unable to appreciate. These opportunities are sometimes apparent due to the visual appearance of the topology, though they are often misleading for the various reasons discussed in chapter 4. It is possible to see occasionally a pair of trunks, running almost parallel for part of their length and a reroute scheme could be devised to allow a sharing of capacity and some cost reduction. It is recognised that there is considerable departure from more usual reduction methods, by proposing visually inspired opportunities, though manual intervention in the design process has been shown to support the idea and has overcome many of the current limitations.

There is considerable scope for extending the TDM core network design system to packet based network technologies such as X.25 and Frame Relay. The major change required of the technique is in the allocation of network resource, trunk capacity, to traffic on its selected paths. In a TDM design system there is a straightforward apportionment of trunk capacity to each selected customer channel, the decision to upgrade to the next available trunk capacity is made when all current capacity is allocated. In a packet design system trunk capacity is not permanently allocated, the critical performance issue is trunk utilization or queuing delay. It is therefore necessary to change the method for determining the need to increase trunk capacity. The packet based trunk selection rule would be based on an estimate of the trunk utilization or delay, when it exceeds a limiting threshold the next available trunk capacity is selected.

APPENDIX A

Core Network Design Results

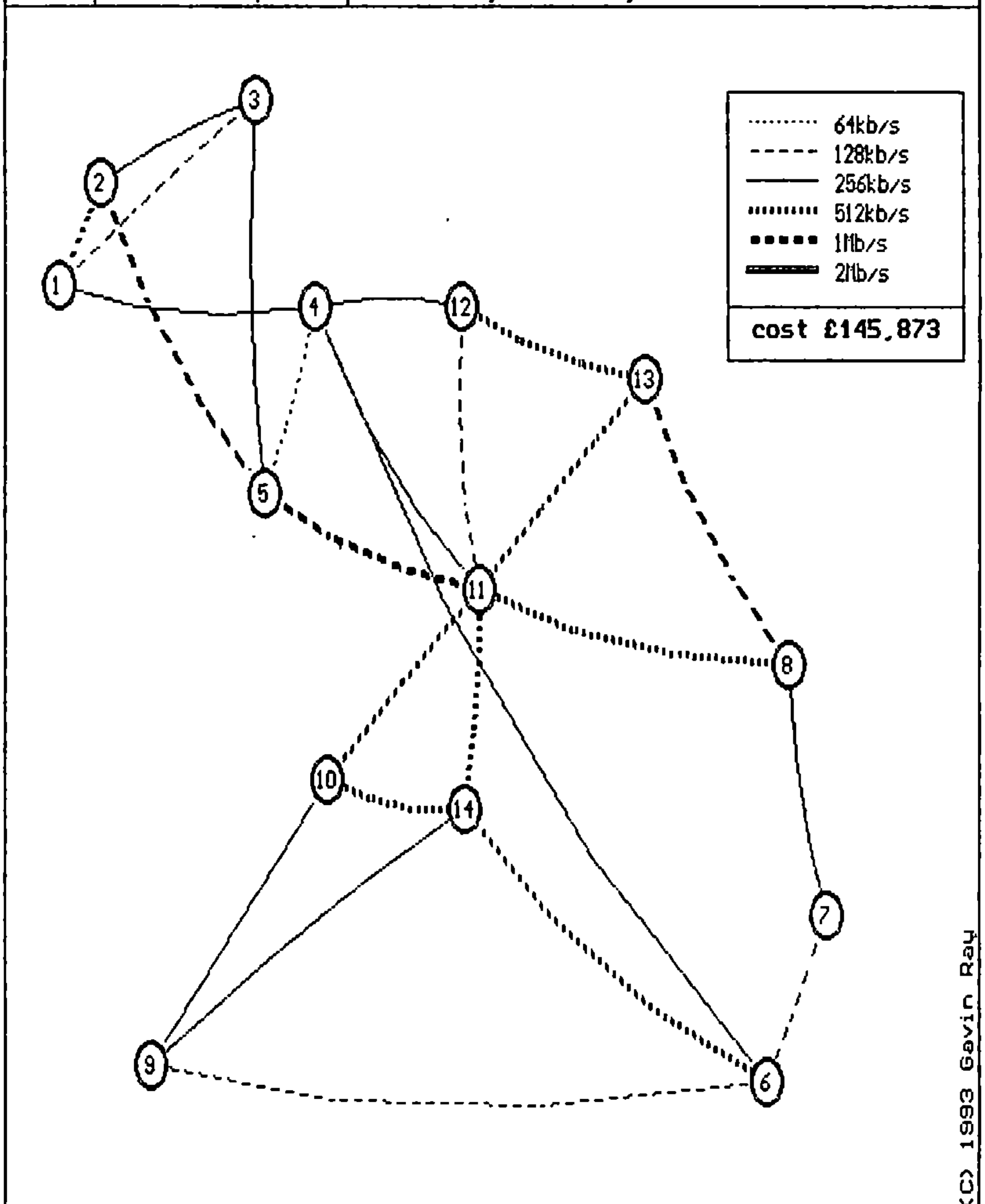
figure A.1



Print Time	=	22:14:08
Optimisation time	=	0 Days, 0h 0m 0s
Traffic Demand File	=	N14.TRF
Link Cost File	=	N14.DAT
Route Definition File	=	JUN29-1.ROU
Link Setup Initial Cost	=	1,000
Link Setup Per km Cost	=	20
Total Network Cost	=	146,291
Cost of Active Capacity	=	125,175
Cost of Spare Capacity	=	21,116
Capital Utilsation	=	85.6 %
Spare / Capacity [channels]	=	22 / 136
Capacity Utilsation	=	83.8 %
Mean No Primary Hops	=	1.88
Mean No Backup Hops	=	2.29
Backup / Primary Ratio	=	1.22
Total Primary Channels	=	66
Total Backup Channels	=	48
Primary/Backup Channel ratio	=	1.38

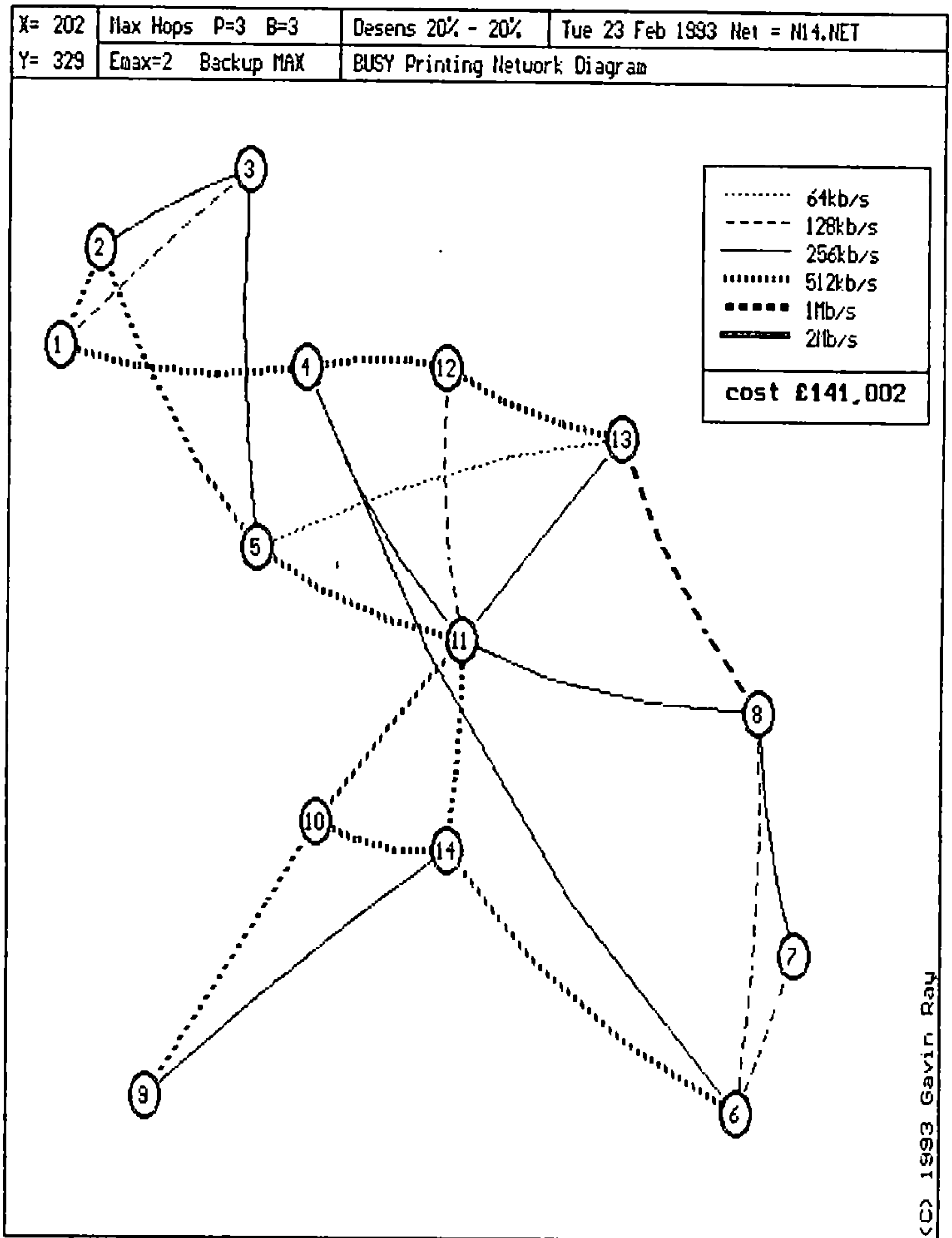
figure A.2

X= 169	Max Hops P=3 B=3	Desens 10% - 10%	Tue 23 Feb 1993 Net = N14.NET
Y= 311	Emax=2 Backup MAX	BUSY Printing Network Diagram	



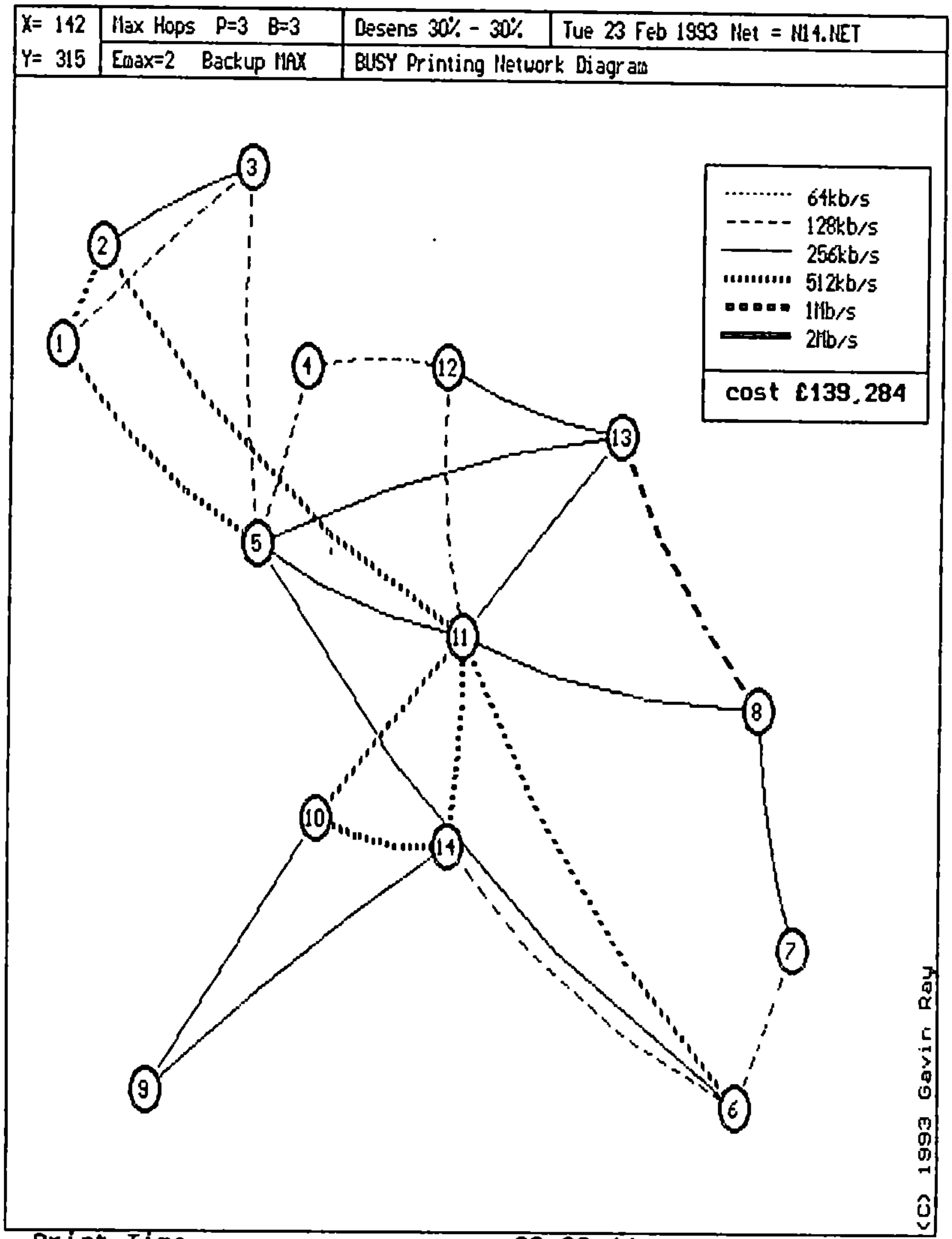
Print Time	=	22:18:18
Optimisation time	=	0 Days, 0h 0m 0s
Traffic Demand File	=	N14.TRF
Link Cost File	=	N14.DAT
Route Definition File	=	JUN29-2.ROU
Link Setup Initial Cost	=	1,000
Link Setup Per km Cost	=	20
Total Network Cost	=	145,873
Cost of Active Capacity	=	115,526
Cost of Spare Capacity	=	30,347
Capital Utilsation	=	79.2 %
Spare / Capacity [channels]	=	39 / 157
Capacity Utilsation	=	75.2 %
Mean No Primary Hops	=	1.88
Mean No Backup Hops	=	2.41
Backup / Primary Ratio	=	1.28
Total Primary Channels	=	67
Total Backup Channels	=	51
Primary/Backup Channel ratio	=	1.31

figure A.3



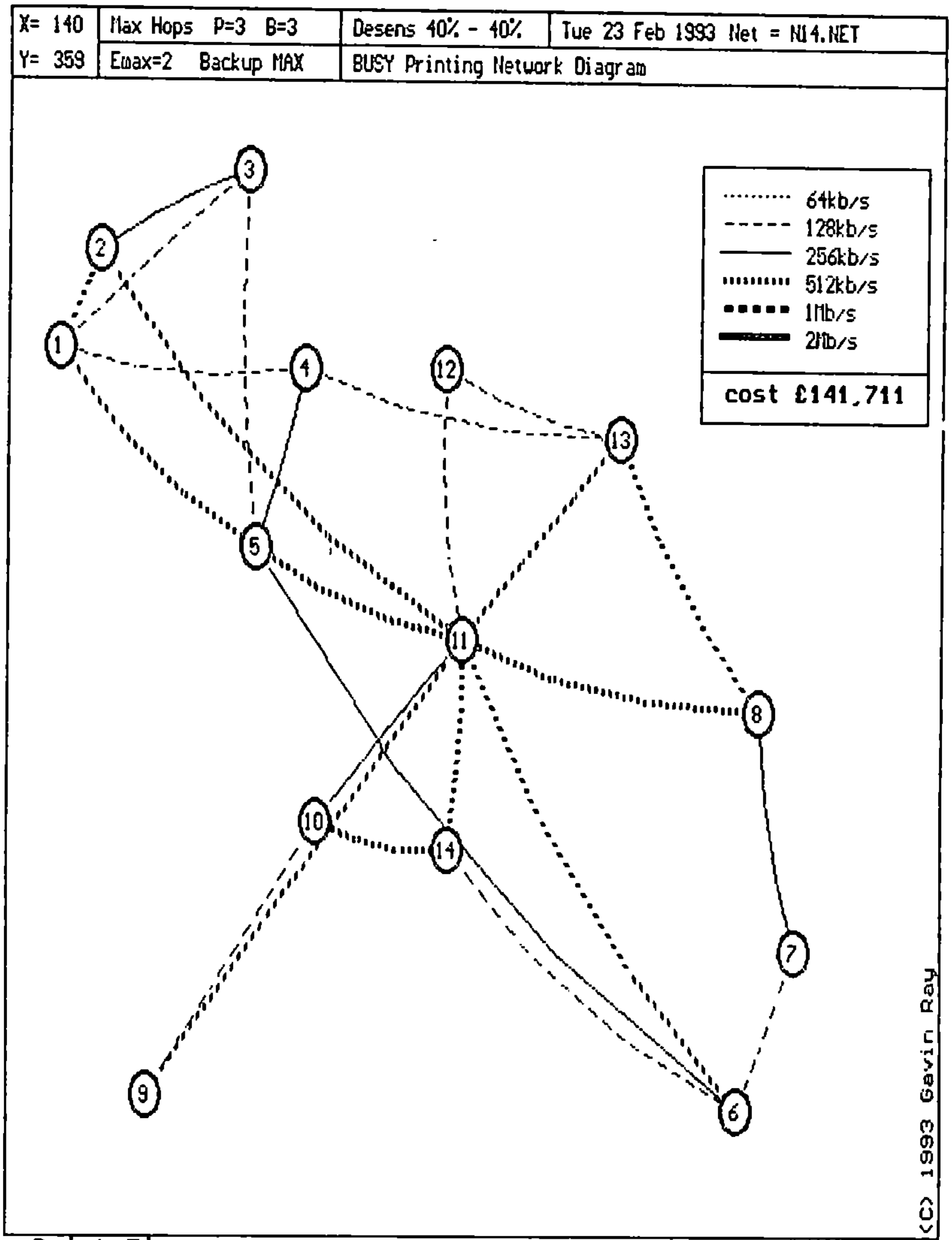
Print Time	=	22:27:14
Optimisation time	=	0 Days, 0h 0m 0s
Traffic Demand File	=	N14.TRF
Link Cost File	=	N14.DAT
Route Definition File	=	JUN29-3.ROU
Link Setup Initial Cost	=	1,000
Link Setup Per km Cost	=	20
Total Network Cost	=	141,002
Cost of Active Capacity	=	117,154
Cost of Spare Capacity	=	23,848
Capital Utilsation	=	83.1 %
Spare / Capacity [channels]	=	29 / 145
Capacity Utilsation	=	80.0 %
Mean No Primary Hops	=	1.94
Mean No Backup Hops	=	2.53
Backup / Primary Ratio	=	1.30
Total Primary Channels	=	67
Total Backup Channels	=	49
Primary/Backup Channel ratio	=	1.37

figure A.4



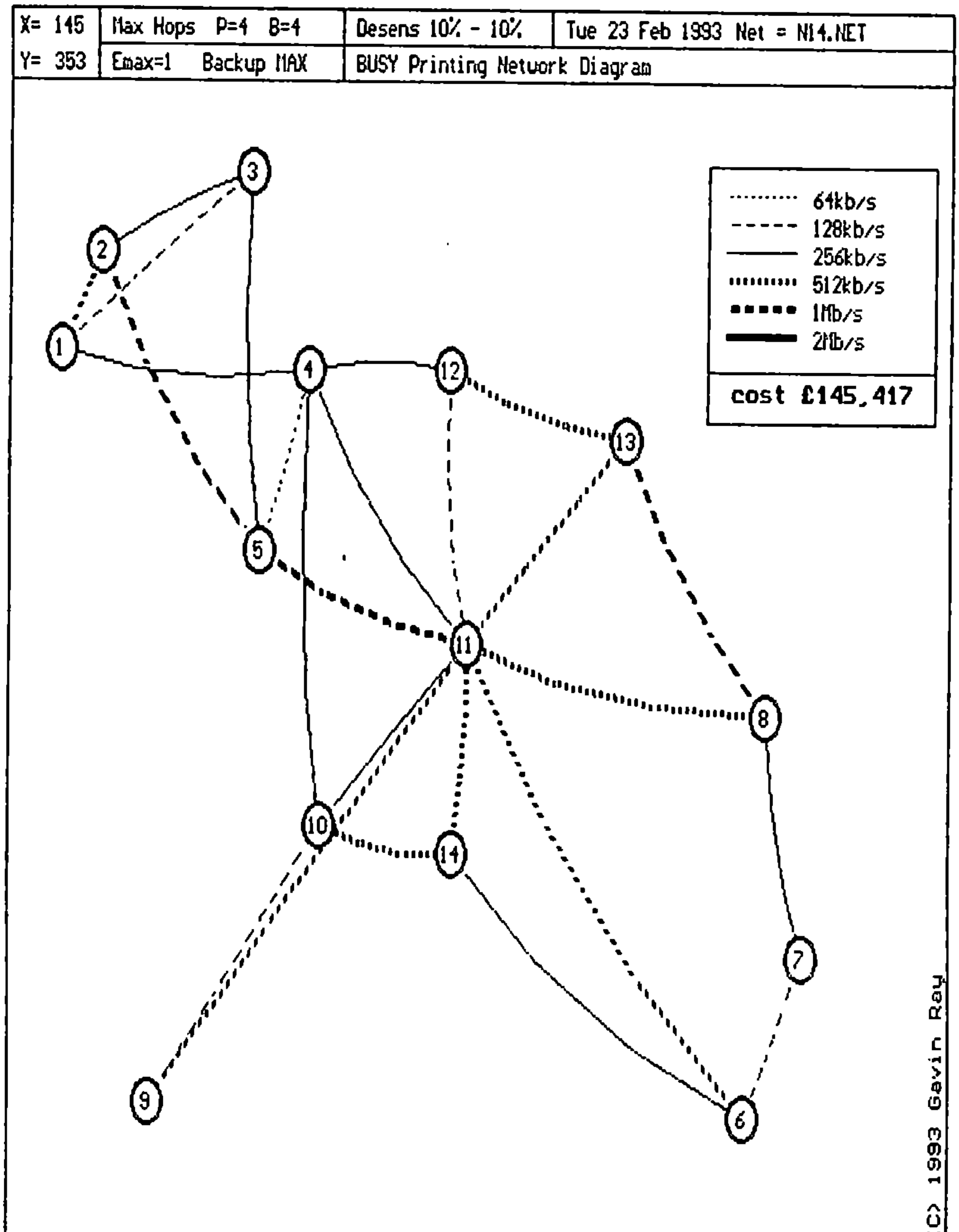
Print Time	=	22:29:44
Optimisation time	=	0 Days, 0h 0m 0s
Traffic Demand File	=	N14.TRF
Link Cost File	=	N14.DAT
Route Definition File	=	JUN29-4.ROU
Link Setup Initial Cost	=	1,000
Link Setup Per km Cost	=	20
Total Network Cost	=	139,284
Cost of Active Capacity	=	118,804
Cost of Spare Capacity	=	20,480
Capital Utilsation	=	85.3 %
Spare / Capacity [channels]	=	25 / 126
Capacity Utilsation	=	80.2 %
Mean No Primary Hops	=	1.59
Mean No Backup Hops	=	2.35
Backup / Primary Ratio	=	1.48
Total Primary Channels	=	55
Total Backup Channels	=	46
Primary/Backup Channel ratio	=	1.20

figure A.5



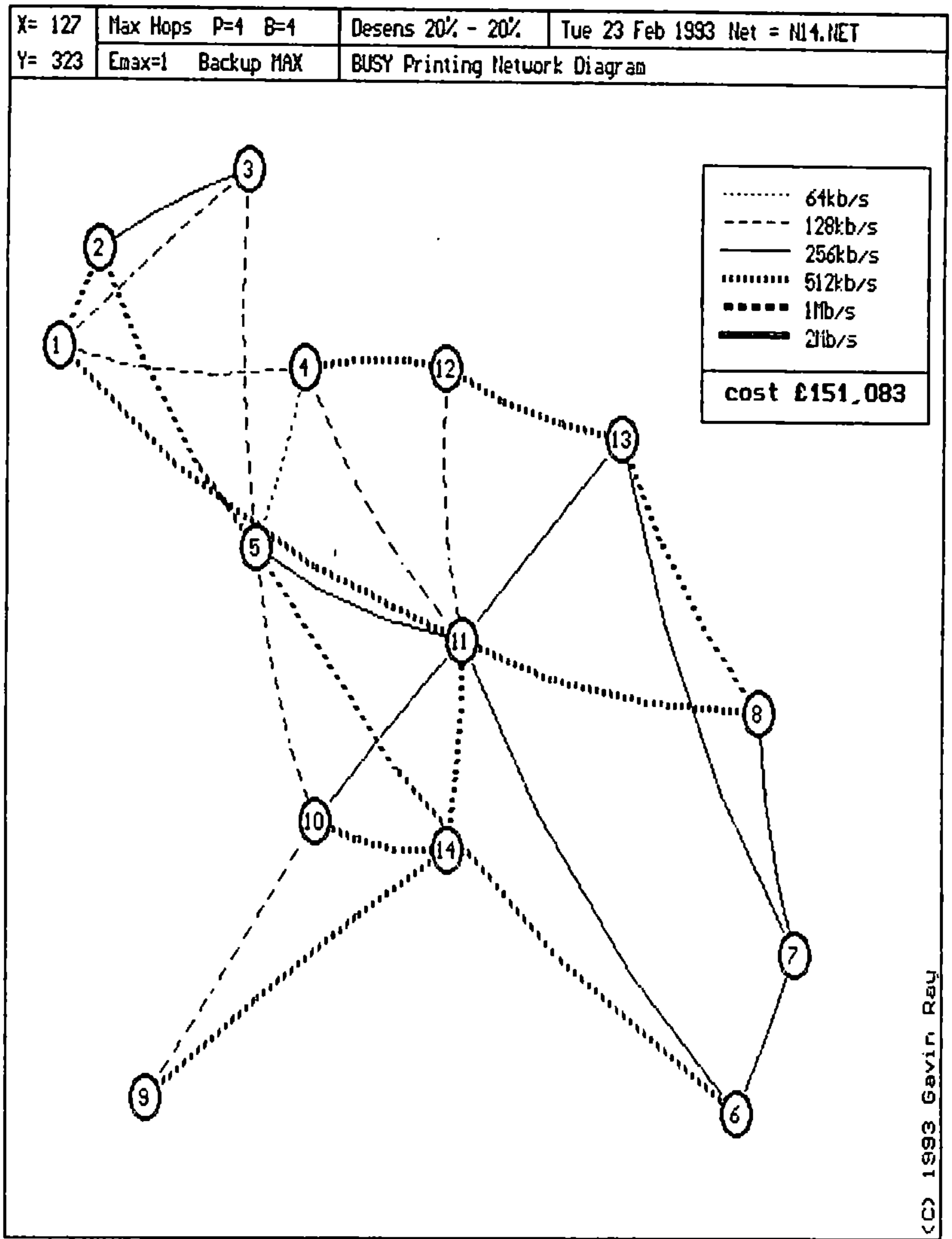
Print Time	=	22:32:40
Optimisation time	=	0 Days, 0h 0m 0s
Traffic Demand File	=	N14.TRF
Link Cost File	=	N14.DAT
Route Definition File	=	JUN29-5.ROU
Link Setup Initial Cost	=	1,000
Link Setup Per km Cost	=	20
Total Network Cost	=	141,711
Cost of Active Capacity	=	119,283
Cost of Spare Capacity	=	22,428
Capital Utilsation	=	84.2 %
Spare / Capacity [channels]	=	24 / 126
Capacity Utilsation	=	81.0 %
Mean No Primary Hops	=	1.59
Mean No Backup Hops	=	2.41
Backup / Primary Ratio	=	1.52
Total Primary Channels	=	53
Total Backup Channels	=	49
Primary/Backup Channel ratio	=	1.08

figure A.6



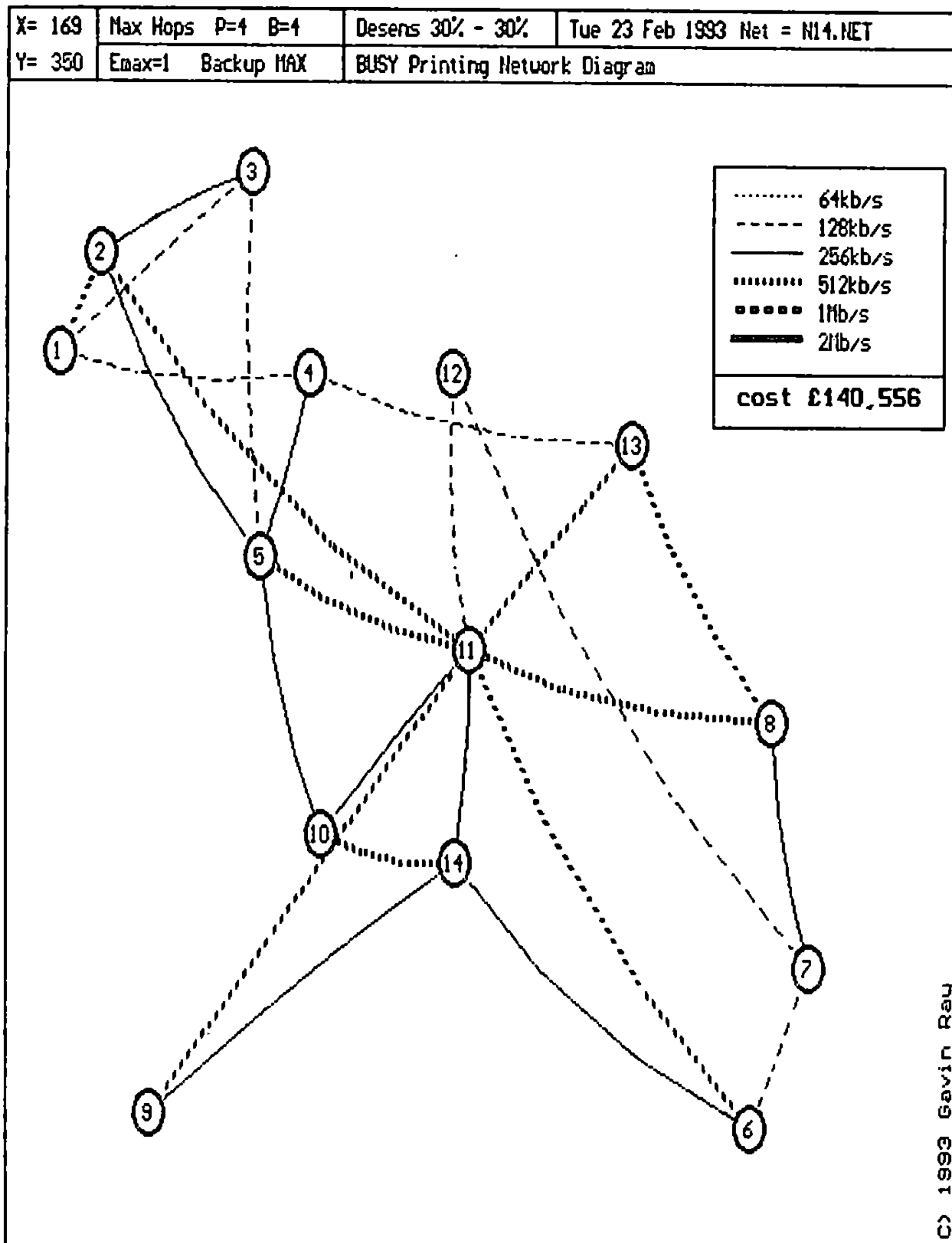

```
Print Time                       = 22:35:18
Optimisation time                = 0 Days, 0h 0m 0s
Traffic Demand File              = N14.TRF
Link Cost File                   = N14.DAT
Route Definition File            = D2SENS1.ROU
Link Setup Initial Cost          = 4,000
Link Setup Per km Cost           = 20
Total Network Cost               = 145,417
Cost of Active Capacity          = 111,583
Cost of Spare Capacity           = 33,834
Capital Utilsation               = 76.7 %
Spare / Capacity [channels]      = 43 / 157
Capacity Utilsation              = 72.6 %
Mean No Primary Hops             = 1.65
Mean No Backup Hops              = 2.41
Backup / Primary Ratio           = 1.46
Total Primary Channels           = 60
Total Backup Channels            = 54
Primary/Backup Channel ratio     = 1.11
```

figure A.7



Print Time	=	22:37:26
Optimisation time	=	0 Days, 0h 0m 0s
Traffic Demand File	=	N14.TRF
Link Cost File	=	N14.DAT
Route Definition File	=	D2SENS2.ROU
Link Setup Initial Cost	=	4,000
Link Setup Per km Cost	=	20
Total Network Cost	=	151,083
Cost of Active Capacity	=	128,813
Cost of Spare Capacity	=	22,270
Capital Utilsation	=	85.3 %
Spare / Capacity [channels]	=	23 / 135
Capacity Utilsation	=	83.0 %
Mean No Primary Hops	=	1.65
Mean No Backup Hops	=	2.12
Backup / Primary Ratio	=	1.29
Total Primary Channels	=	59
Total Backup Channels	=	53
Primary/Backup Channel ratio	=	1.11

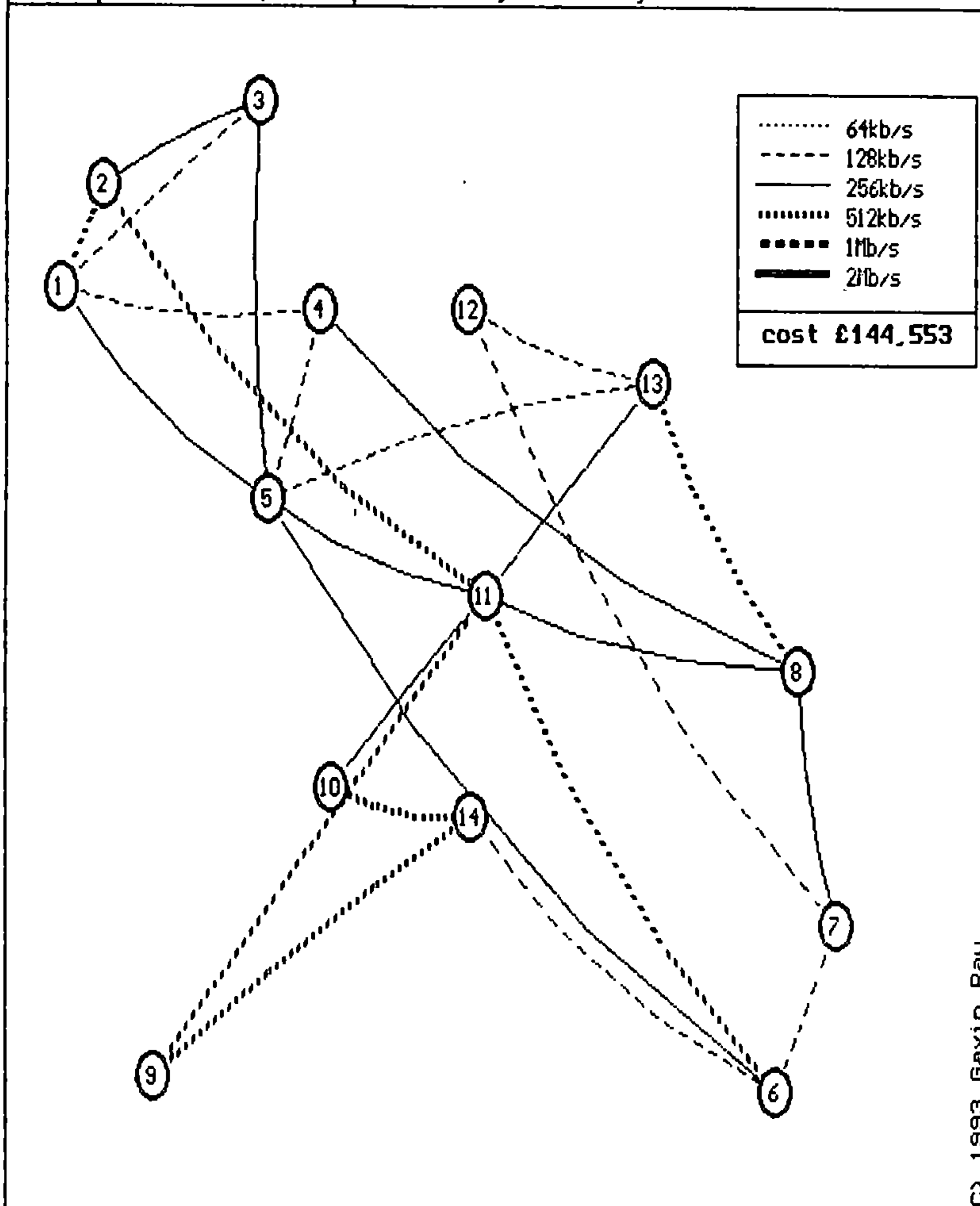
figure A.8




```
Print Time                     = 22:39:34
Optimisation time              = 0 Days, 0h 0m 0s
Traffic Demand File            = N14.TRF
Link Cost File                 = N14.DAT
Route Definition File          = D2SENS3.ROU
Link Setup Initial Cost        = 4,000
Link Setup Per km Cost         = 20
Total Network Cost             = 140,556
Cost of Active Capacity        = 118,404
Cost of Spare Capacity         = 22,153
Capital Utilsation             = 84.2 %
Spare / Capacity [channels]    = 23 / 122
Capacity Utilsation            = 81.1 %
Mean No Primary Hops           = 1.29
Mean No Backup Hops            = 2.12
Backup / Primary Ratio         = 1.64
Total Primary Channels         = 47
Total Backup Channels          = 52
Primary/Backup Channel ratio   = 0.90
```

figure A.9

X= 163	Max Hops P=4 B=4	Desens 40% - 40%	Tue 23 Feb 1993 Net = N14.NET
Y= 331	Emax=1 Backup MAX	BUSY Printing Network Diagram	




```
Print Time                    = 22:41:40
Optimisation time             = 0 Days, 0h 0m 0s
Traffic Demand File           = N14.TRF
Link Cost File                = N14.DAT
Route Definition File         = D2SENS4.ROU
Link Setup Initial Cost       = 4,000
Link Setup Per km Cost        = 20
Total Network Cost            = 144,553
Cost of Active Capacity       = 126,387
Cost of Spare Capacity        = 18,166
Capital Utilsation            = 87.4 %
Spare / Capacity [channels]   = 18 / 112
Capacity Utilsation           = 83.9 %
Mean No Primary Hops          = 1.06
Mean No Backup Hops           = 2.47
Backup / Primary Ratio        = 2.33
Total Primary Channels        = 40
Total Backup Channels         = 54
Primary/Backup Channel ratio  = 0.74
```

APPENDIX B

Core Network Design Report

figure B.1 Text Report for Topology D2SENS1

```
(C) 1993  Gavin Ray - AT&T ISTEL      The Network Optimisation Engine

Version : 22.1.93      Jan 1993

Session by           : Gavin Ray  (the Author)
Todays Date          : Mon 22 Feb 1993
Current Time         : 20:56:19
Session Duration     : 0 Days, 0h 0m 11s

Network Model Family  : New
Network Name          : N14.NET
Demand Traffic File   : N14.TRF
Link Cost File        : N14.DAT
Route Definition File : D2SENS1.ROU

Primary Hop limit      = 4
Backup  Hop limit      = 4

Max Elliptic Search Nodes     = 1 + Max Hops per route
Costing Desensitivity MAX     = 10%
Costing Desensitivity MIN     = 10%
Link Setup Cost               = 4,000
Link Setup Cost (per km)      = 20
Cost Inequality Limit         = 0.8

DESIGN SUMMARY

Total Network Cost            = 145,417
Cost of Spare Capacity        = 33,834
Capacity Utilsation           = 76.7%
Mean No Primary Hops          = 1.81

NODE POSITIONS :

A          Node  1   ---    X = 130     Y = 132
B          Node  2   ---    X = 153     Y =  94
C          Node  3   ---    X = 239     Y =  63
D          Node  4   ---    X = 272     Y = 140
E          Node  5   ---    X = 244     Y = 210
F          Node  6   ---    X = 522     Y = 430
G          Node  7   ---    X = 555     Y = 368
H          Node  8   ---    X = 534     Y = 274
I          Node  9   ---    X = 181     Y = 424
J          Node 10   ---    X = 279     Y = 317
K          Node 11   ---    X = 363     Y = 246
L          Node 12   ---    X = 353     Y = 140
M          Node 13   ---    X = 454     Y = 167
N          Node 14   ---    X = 355     Y = 328

NUMERICAL ORDER OF TRUNKS
```

TRUNK	Source	Destination	Capacity	Cost	Primary	Backup	Spare	
1	2-B	1-A	512kb/s	4,190	4	3	[1]	£524
2	3-C	1-A	128kb/s	4,161	0	2	[0]	£0
3	3-C	2-B	256kb/s	4,440	1	2	[1]	£1110
4	4-D	1-A	256kb/s	5,240	0	3	[1]	£1310
5	5-E	2-B	1Mb/s	9,517	8	2	[6]	£3569
6	5-E	3-C	256kb/s	5,317	2	1	[1]	£1329
7	5-E	4-D	64kb/s	3,377	0	1	[0]	£0
8	7-G	6-F	128kb/s	3,632	0	2	[0]	£0
9	8-H	7-G	256kb/s	4,517	4	0	[0]	£0

10	10-J	4-D	256kb/s	5,790	0	3	[1]	£1448
11	10-J	9-I	128kb/s	4,306	0	2	[0]	£0
12	11-K	4-D	256kb/s	5,200	0	3	[1]	£1300
13	11-K	5-E	1Mb/s	8,495	9	0	[7]	£3717
14	11-K	6-F	512kb/s	9,517	5	2	[1]	£1190
15	11-K	8-H	512kb/s	7,644	2	3	[3]	£2867
16	11-K	9-I	512kb/s	9,823	5	0	[3]	£3684
17	11-K	10-J	256kb/s	4,732	0	4	[0]	£0
18	12-L	4-D	256kb/s	4,276	2	2	[0]	£0
19	12-L	11-K	128kb/s	3,958	0	2	[0]	£0
20	13-M	8-H	1Mb/s	8,905	7	2	[7]	£3896
21	13-M	11-K	512kb/s	6,230	2	3	[3]	£2336
22	13-M	12-L	512kb/s	5,802	4	2	[2]	£1451
23	14-N	6-F	256kb/s	6,082	0	3	[1]	£1521
24	14-N	10-J	512kb/s	5,058	4	3	[1]	£632
25	14-N	11-K	512kb/s	5,208	1	4	[3]	£1953

Total Network Cost = 145,417 Total Spare Cost = 33,834

LINK TRAFFIC DEMAND (Channels)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	A	--	2	0	0	0	0	0	2	0	0	0	0	0	0
2	B	2	--	1	0	0	3	0	0	0	0	3	0	0	0
3	C	0	1	--	0	2	0	0	0	0	0	0	0	0	0
4	D	0	0	0	--	0	0	0	2	0	0	0	0	0	0
5	E	0	0	2	0	--	0	0	0	0	0	0	0	1	0
6	F	0	3	0	0	0	--	0	0	2	0	0	0	0	0
7	G	0	0	0	0	0	0	--	2	0	0	0	2	0	0
8	H	2	0	0	2	0	0	2	--	0	0	0	0	3	0
9	I	0	0	0	0	0	2	0	0	--	0	2	0	0	1
10	J	0	0	0	0	0	0	0	0	0	--	0	0	0	4
11	K	0	3	0	0	0	0	0	0	2	0	--	0	1	1
12	L	0	0	0	0	0	0	2	0	0	0	0	--	0	0
13	M	0	0	0	0	1	0	0	3	0	0	1	0	--	0
14	N	0	0	0	0	0	0	0	0	1	4	1	0	0	--

LINK TRAFFIC ALLOCATION (Channels)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	A	--	7	2	3	0	0	0	0	0	0	0	0	0	0
2	B	7	--	3	0	10	0	0	0	0	0	0	0	0	0
3	C	2	3	--	0	3	0	0	0	0	0	0	0	0	0
4	D	3	0	0	--	1	0	0	0	0	3	3	4	0	0
5	E	0	10	3	1	--	0	0	0	0	0	9	0	0	0
6	F	0	0	0	0	0	--	2	0	0	0	7	0	0	3
7	G	0	0	0	0	0	2	--	4	0	0	0	0	0	0
8	H	0	0	0	0	0	0	4	--	0	0	5	0	9	0
9	I	0	0	0	0	0	0	0	0	--	2	5	0	0	0
10	J	0	0	0	3	0	0	0	0	2	--	4	0	0	7

11	K	0	0	0	3	9	7	0	5	5	4	--	2	5	5
12	L	0	0	0	4	0	0	0	0	0	0	2	--	6	0
13	M	0	0	0	0	0	0	0	9	0	0	5	6	--	0
14	N	0	0	0	0	0	3	0	0	0	7	5	0	0	--

Route Allocations - by NODE

Routes from node 1
1100: Primary = 2 1 cost = 1,048 direct cost = 3,400
ratio = 0.31
1100: Backup = 2 3 1 cost = 1,110 direct cost = 3,400
ratio = 0.33
1100: Primary = 8 11 5 2 1 cost = 1,048 direct cost = 6,854
ratio = 0.15
1100: Backup = 8 13 12 4 1 cost = 1,670 direct cost = 6,854
ratio = 0.24

Routes from node 2
1100: Primary = 2 1 cost = 1,048 direct cost = 3,400
ratio = 0.31
1100: Backup = 2 3 1 cost = 1,110 direct cost = 3,400
ratio = 0.33
1100: Primary = 3 2 cost = 1,110 direct cost = 3,457
ratio = 0.32
1100: Backup = 3 5 2 cost = 2,659 direct cost = 3,457
ratio = 0.77
1100: Primary = 6 11 5 2 cost = 0 direct cost = 8,145
ratio = 0.00
1100: Backup = 6 14 10 4 1 2 cost = 3,576 direct cost = 8,145
ratio = 0.44
1100: Primary = 11 5 2 cost = 0 direct cost = 5,312
ratio = 0.00
1100: Backup = 11 4 1 2 cost = 1,048 direct cost = 5,312
ratio = 0.20

Routes from node 3
1100: Primary = 3 2 cost = 1,110 direct cost = 3,457
ratio = 0.32
1100: Backup = 3 5 2 cost = 2,659 direct cost = 3,457
ratio = 0.77
1100: Primary = 5 3 cost = 2,659 direct cost = 4,324
ratio = 0.61
1100: Backup = 5 2 3 cost = 1,110 direct cost = 4,324
ratio = 0.26

Routes from node 4
1100: Primary = 8 13 12 4 cost = 1,670 direct cost = 5,649
ratio = 0.30
1100: Backup = 8 11 4 cost = 0 direct cost = 5,649
ratio = 0.00

Routes from node 5
1100: Primary = 5 3 cost = 2,659 direct cost = 4,324
ratio = 0.61
1100: Backup = 5 2 3 cost = 1,110 direct cost = 4,324
ratio = 0.26
1100: Primary = 13 11 5 cost = 779 direct cost = 4,072
ratio = 0.19
1100: Backup = 13 12 4 5 cost = 0 direct cost = 4,072
ratio = 0.00

Routes from node 6
1100: Primary = 6 11 5 2 cost = 0 direct cost = 8,145
ratio = 0.00
1100: Backup = 6 14 10 4 1 2 cost = 3,576 direct cost = 8,145
ratio = 0.44
1100: Primary = 9 11 6 cost = 2,456 direct cost = 6,069
ratio = 0.40
1100: Backup = 9 10 14 6 cost = 2,529 direct cost = 6,069
ratio = 0.42

```
Routes from node 7
1100:  Primary = 8 7                    cost = 2,258    direct cost = 3,867
ratio = 0.58
1100:   Backup = 8 11 6 7               cost = 0        direct cost = 3,867
ratio = 0.00
1100:  Primary = 12 13 8 7              cost = 3,928    direct cost = 5,742
ratio = 0.68
1100:   Backup = 12 11 6 7              cost = 0        direct cost = 5,742
ratio = 0.00

Routes from node 8
1100:  Primary = 8 11 5 2 1             cost = 1,048    direct cost = 6,854
ratio = 0.15
1100:   Backup = 8 13 12 4 1            cost = 1,670    direct cost = 6,854
ratio = 0.24
1100:  Primary = 8 13 12 4              cost = 1,670    direct cost = 5,649
ratio = 0.30
1100:   Backup = 8 11 4                 cost = 0        direct cost = 5,649
ratio = 0.00
1100:  Primary = 8 7                    cost = 2,258    direct cost = 3,867
ratio = 0.58
1100:   Backup = 8 11 6 7               cost = 0        direct cost = 3,867
ratio = 0.00
1100:  Primary = 13 8                   cost = 1,670    direct cost = 3,828
ratio = 0.44
1100:   Backup = 13 11 8                cost = 779      direct cost = 3,828
ratio = 0.20

Routes from node 9
1100:  Primary = 9 11 6                 cost = 2,456    direct cost = 6,069
ratio = 0.40
1100:   Backup = 9 10 14 6              cost = 2,529    direct cost = 6,069
ratio = 0.42
1100:  Primary = 11 9                   cost = 2,456    direct cost = 5,291
ratio = 0.46
1100:   Backup = 11 10 9                cost = 0        direct cost = 5,291
ratio = 0.00
1100:  Primary = 14 11 9                cost = 3,107    direct cost = 3,994
ratio = 0.78
1100:   Backup = 14 10 9                cost = 2,529    direct cost = 3,994
ratio = 0.63

Routes from node 10
1100:  Primary = 14 10                  cost = 2,529    direct cost = 4,209
ratio = 0.60
1100:   Backup = 14 11 10               cost = 651      direct cost = 4,209
ratio = 0.15

Routes from node 11
1100:  Primary = 11 5 2                 cost = 0        direct cost = 5,312
ratio = 0.00
1100:   Backup = 11 4 1 2               cost = 1,048    direct cost = 5,312
ratio = 0.20
1100:  Primary = 11 9                   cost = 2,456    direct cost = 5,291
ratio = 0.46
1100:   Backup = 11 10 9                cost = 0        direct cost = 5,291
ratio = 0.00
1100:  Primary = 13 11                  cost = 779      direct cost = 3,603
ratio = 0.22
1100:   Backup = 13 8 11                cost = 1,670    direct cost = 3,603
ratio = 0.46

Routes from node 12
1100:  Primary = 12 13 8 7              cost = 3,928    direct cost = 5,742
ratio = 0.68
1100:   Backup = 12 11 6 7              cost = 0        direct cost = 5,742
ratio = 0.00

Routes from node 13
1100:  Primary = 13 11 5                cost = 779      direct cost = 4,072
ratio = 0.19
1100:   Backup = 13 12 4 5              cost = 0        direct cost = 4,072
ratio = 0.00
1100:  Primary = 13 8                   cost = 1,670    direct cost = 3,828
ratio = 0.44
1100:   Backup = 13 11 8                cost = 779      direct cost = 3,828
```

```
ratio = 0.20
1100:  Primary = 13 11                        cost = 779       direct cost = 3,603
ratio = 0.22
1100:   Backup = 13 8 11                      cost = 1,670     direct cost = 3,603
ratio = 0.46

Routes from node 14
1100:  Primary = 14 11 9                      cost = 3,107     direct cost = 3,994
ratio = 0.78
1100:   Backup = 14 10 9                      cost = 2,529     direct cost = 3,994
ratio = 0.63
1100:  Primary = 14 10                        cost = 2,529     direct cost = 4,209
ratio = 0.60
1100:   Backup = 14 11 10                     cost = 651       direct cost = 4,209
ratio = 0.15
```

Route Allocations   - by LINK

```
Routes Using Trunk 2 to 1
   Link Capacity Allocated    = 512kb/s
   Link Cost (Per Channel)    = 524
   Primary Channels Allocated = 4
   Backup  Channels Allocated = 3
     Primary = 2:1                        Traffic = 2 ch
     Primary = 8:11:5:2:1                 Traffic = 2 ch
     Backup  = 6:14:10:4:1:2              Traffic = 3 ch
     Backup  = 11:4:1:2                   Traffic = 3 ch

Routes Using Trunk 3 to 1
   Link Capacity Allocated    = 128kb/s
   Link Cost (Per Channel)    = 2,081
   Primary Channels Allocated = 0
   Backup  Channels Allocated = 2
     Backup  = 2:3:1                     Traffic = 2 ch

Routes Using Trunk 3 to 2
   Link Capacity Allocated    = 256kb/s
   Link Cost (Per Channel)    = 1,110
   Primary Channels Allocated = 1
   Backup  Channels Allocated = 2
     Primary = 3:2                       Traffic = 1 ch
     Backup  = 2:3:1                     Traffic = 2 ch
     Backup  = 5:2:3                     Traffic = 2 ch

Routes Using Trunk 4 to 1
   Link Capacity Allocated    = 256kb/s
   Link Cost (Per Channel)    = 1,310
   Primary Channels Allocated = 0
   Backup  Channels Allocated = 3
     Backup  = 6:14:10:4:1:2             Traffic = 3 ch
     Backup  = 8:13:12:4:1               Traffic = 2 ch
     Backup  = 11:4:1:2                  Traffic = 3 ch

Routes Using Trunk 5 to 2
   Link Capacity Allocated    =   1Mb/s
   Link Cost (Per Channel)    = 595
   Primary Channels Allocated = 8
   Backup  Channels Allocated = 2
     Primary = 6:11:5:2                  Traffic = 3 ch
     Primary = 8:11:5:2:1                Traffic = 2 ch
     Primary = 11:5:2                    Traffic = 3 ch
     Backup  = 3:5:2                     Traffic = 1 ch
     Backup  = 5:2:3                     Traffic = 2 ch

Routes Using Trunk 5 to 3
   Link Capacity Allocated    = 256kb/s
   Link Cost (Per Channel)    = 1,329
   Primary Channels Allocated = 2
   Backup  Channels Allocated = 1
     Primary = 5:3                       Traffic = 2 ch
     Backup  = 3:5:2                     Traffic = 1 ch

Routes Using Trunk 5 to 4
   Link Capacity Allocated    =  64kb/s
   Link Cost (Per Channel)    = 3,377
```

```
   Primary Channels Allocated = 0
   Backup  Channels Allocated = 1
     Backup  = 13:12:4:5                        Traffic = 1 ch

Routes Using Trunk 7 to 6
   Link Capacity Allocated    = 128kb/s
   Link Cost (Per Channel)    = 1,816
   Primary Channels Allocated = 0
   Backup  Channels Allocated = 2
     Backup  = 8:11:6:7                         Traffic = 2 ch
     Backup  = 12:11:6:7                        Traffic = 2 ch

Routes Using Trunk 8 to 7
   Link Capacity Allocated    = 256kb/s
   Link Cost (Per Channel)    = 1,129
   Primary Channels Allocated = 4
   Backup  Channels Allocated = 0
     Primary = 8:7                              Traffic = 2 ch
     Primary = 12:13:8:7                        Traffic = 2 ch

Routes Using Trunk 10 to 4
   Link Capacity Allocated    = 256kb/s
   Link Cost (Per Channel)    = 1,448
   Primary Channels Allocated = 0
   Backup  Channels Allocated = 3
     Backup  = 6:14:10:4:1:2                    Traffic = 3 ch

Routes Using Trunk 10 to 9
   Link Capacity Allocated    = 128kb/s
   Link Cost (Per Channel)    = 2,153
   Primary Channels Allocated = 0
   Backup  Channels Allocated = 2
     Backup  = 9:10:14:6                        Traffic = 2 ch
     Backup  = 11:10:9                          Traffic = 2 ch
     Backup  = 14:10:9                          Traffic = 1 ch

Routes Using Trunk 11 to 4
   Link Capacity Allocated    = 256kb/s
   Link Cost (Per Channel)    = 1,300
   Primary Channels Allocated = 0
   Backup  Channels Allocated = 3
     Backup  = 8:11:4                           Traffic = 2 ch
     Backup  = 11:4:1:2                         Traffic = 3 ch

Routes Using Trunk 11 to 5
   Link Capacity Allocated    =   1Mb/s
   Link Cost (Per Channel)    = 531
   Primary Channels Allocated = 9
   Backup  Channels Allocated = 0
     Primary = 6:11:5:2                         Traffic = 3 ch
     Primary = 8:11:5:2:1                       Traffic = 2 ch
     Primary = 11:5:2                           Traffic = 3 ch
     Primary = 13:11:5                          Traffic = 1 ch

Routes Using Trunk 11 to 6
   Link Capacity Allocated    = 512kb/s
   Link Cost (Per Channel)    = 1,190
   Primary Channels Allocated = 5
   Backup  Channels Allocated = 2
     Primary = 6:11:5:2                         Traffic = 3 ch
     Primary = 9:11:6                           Traffic = 2 ch
     Backup  = 8:11:6:7                         Traffic = 2 ch
     Backup  = 12:11:6:7                        Traffic = 2 ch

Routes Using Trunk 11 to 8
   Link Capacity Allocated    = 512kb/s
   Link Cost (Per Channel)    = 956
   Primary Channels Allocated = 2
   Backup  Channels Allocated = 3
     Primary = 8:11:5:2:1                       Traffic = 2 ch
     Backup  = 8:11:4                           Traffic = 2 ch
     Backup  = 8:11:6:7                         Traffic = 2 ch
     Backup  = 13:11:8                          Traffic = 3 ch
     Backup  = 13:8:11                          Traffic = 1 ch

Routes Using Trunk 11 to 9
```

```
    Link Capacity Allocated    = 512kb/s
    Link Cost (Per Channel)    = 1,228
    Primary Channels Allocated = 5
    Backup  Channels Allocated = 0
      Primary = 9:11:6                          Traffic = 2 ch
      Primary = 11:9                            Traffic = 2 ch
      Primary = 14:11:9                         Traffic = 1 ch

Routes Using Trunk 11 to 10
    Link Capacity Allocated    = 256kb/s
    Link Cost (Per Channel)    = 1,183
    Primary Channels Allocated = 0
    Backup  Channels Allocated = 4
      Backup  = 11:10:9                         Traffic = 2 ch
      Backup  = 14:11:10                        Traffic = 4 ch

Routes Using Trunk 12 to 4
    Link Capacity Allocated    = 256kb/s
    Link Cost (Per Channel)    = 1,069
    Primary Channels Allocated = 2
    Backup  Channels Allocated = 2
      Primary = 8:13:12:4                       Traffic = 2 ch
      Backup  = 8:13:12:4:1                     Traffic = 2 ch
      Backup  = 13:12:4:5                       Traffic = 1 ch

Routes Using Trunk 12 to 11
    Link Capacity Allocated    = 128kb/s
    Link Cost (Per Channel)    = 1,979
    Primary Channels Allocated = 0
    Backup  Channels Allocated = 2
      Backup  = 12:11:6:7                       Traffic = 2 ch

Routes Using Trunk 13 to 8
    Link Capacity Allocated    =   1Mb/s
    Link Cost (Per Channel)    = 557
    Primary Channels Allocated = 7
    Backup  Channels Allocated = 2
      Primary = 8:13:12:4                       Traffic = 2 ch
      Primary = 12:13:8:7                       Traffic = 2 ch
      Primary = 13:8                            Traffic = 3 ch
      Backup  = 8:13:12:4:1                     Traffic = 2 ch
      Backup  = 13:8:11                         Traffic = 1 ch

Routes Using Trunk 13 to 11
    Link Capacity Allocated    = 512kb/s
    Link Cost (Per Channel)    = 779
    Primary Channels Allocated = 2
    Backup  Channels Allocated = 3
      Primary = 13:11:5                         Traffic = 1 ch
      Primary = 13:11                           Traffic = 1 ch
      Backup  = 13:11:8                         Traffic = 3 ch

Routes Using Trunk 13 to 12
    Link Capacity Allocated    = 512kb/s
    Link Cost (Per Channel)    = 725
    Primary Channels Allocated = 4
    Backup  Channels Allocated = 2
      Primary = 8:13:12:4                       Traffic = 2 ch
      Primary = 12:13:8:7                       Traffic = 2 ch
      Backup  = 8:13:12:4:1                     Traffic = 2 ch
      Backup  = 13:12:4:5                       Traffic = 1 ch

Routes Using Trunk 14 to 6
    Link Capacity Allocated    = 256kb/s
    Link Cost (Per Channel)    = 1,521
    Primary Channels Allocated = 0
    Backup  Channels Allocated = 3
      Backup  = 6:14:10:4:1:2                   Traffic = 3 ch
      Backup  = 9:10:14:6                       Traffic = 2 ch

Routes Using Trunk 14 to 10
    Link Capacity Allocated    = 512kb/s
    Link Cost (Per Channel)    = 632
    Primary Channels Allocated = 4
    Backup  Channels Allocated = 3
      Primary = 14:10                           Traffic = 4 ch
```

```
    Backup  = 6:14:10:4:1:2                       Traffic = 3 ch
    Backup  = 9:10:14:6                           Traffic = 2 ch
    Backup  = 14:10:9                             Traffic = 1 ch

Routes Using Trunk 14 to 11
   Link Capacity Allocated    = 512kb/s
   Link Cost (Per Channel)    = 651
   Primary Channels Allocated = 1
   Backup  Channels Allocated = 4
     Primary = 14:11:9                            Traffic = 1 ch
     Backup  = 14:11:10                           Traffic = 4 ch
```

End of Report File

REFERENCES

Agarwal Y.K.

'An Algorithm for Designing Survivable Networks', AT&T Technical Journal May/June 1989.

Akimaru H, Kuribayashi H. & Inoue T.

'Approximate Evaluation for Mixed Delay and Loss Systems with Renewal and Poisson Inputs', IEEE Trans on Communications, vol 36, no 7, July 1988.

Bahl L.R. & Tang D.T.

'Optimisation of Concentrator Locations in Teleprocessing Networks', Symposium on Computer Communication Networks and Teletraffic, Polytechnic Institute of Brooklyn, April 1972.

Balakrishnan A. & Graves S.C.

'A Composite Algorithm for a Concave-Cost Network Flow Problem', Networks, vol 19, 1989.

Boorstyn R.R. & Frank H.

'Large-Scale Network Topological Optimisation', IEEE Trans Communications vol COM-25 No 1 Jan 1977.

Boyce D.E, Fahri A. & Weischal R.

'Optimal Network Problem: A Branch and Bound Algorithm', Environment and Planning, 1973, vol 5.

Cheng Y-C, Houck D.J. Liu J-M, Meketon M.S, Slutsman L, Vanderbei R.J and Wang P.

'The AT&T Korbx System', AT&T Technical Journal May/June 1989.

Cox T, Dix F, Hemrick C & McRoberts J.

'SMDS: The Beginning of WAN Super-highways', Data Communications, Apr 1991.

Danzig G.B, Harvey R.P, Lansdowne Z.F, Robinson D.W. & Maier S.F.

'Formulating and Solving the Network Design Problem by Decomposition', Transpn Res-B vol 13B pp5-17, Pergamon Press 1979.

Deagan J.J., Luderer G.W.R. & Vaidya A.K.

'Fast Packet Technology for Future Switches', AT&T Technical Journal, Mar 1989.

Dettra, R.A.

'Metropolitan Area Networks - Data Transport Highways', AT&T Technology, volume 6, number 1. 1991.

Emmett A.

'Karmarker's Algorithm: a threat to Simplex?' IEEE Spectrum Dec 1985.

Esfahanian A. & Hakim S.L.

'Fault-Tolerant Routing in DeBruijn Communication Networks', IEEE Trans on Computers, vol 34, No 9, 1985.

Frank H. & Chou W.

'Routing in Computer Networks', Networks, 1 99-112, John Wiley and Sons Inc, 1971.

Frank H. & Chou W.

'Topological Optimisation of Computer Networks', Proc IEEE vol 60 p1385-1397 Nov 1972.

Frank H. & Chou W.

'Network Properties of the ARPA Computer Network', Networks, vol 4, John Wiley & Sons, 1974.

Frank H, Frisch I.T. & Chou W.

'Topological Considerations in the Design of the ARPA Computer Network', 1970.

Frank H, Frisch.T, van Slyke R. & Chou W.S.

'Optimal Design of Centralised Computer Networks', Networks, vol 1, 1971.

Frank H, Kahn R.E. & Kleinrock L.

'Computer Communication Network Design - Experience with Theory and Practice', Proc Spring Joint Computer Conference, USA. 1972.

Gasparo D. M.

'Public Frame Relay: Price Cutting Isnt Enough', Data Communications, April 1992

Gavish B. & Neuman I.

'A System for Routing and Capacity Assignment in Computer Communication Networks', IEEE Trans on Communications vol 37 no 4 April 1989.

Gerla M.

'Deterministic and Adaptive Routing Policies in Packet Switched Computer Networks', Proc ACM-IEEE 3rd Data Commun Symp Tampa, Florida, Nov 13-15 1973.

Gerla M & Kleinrock L.

'On the Topological Design of Distributed Computer Networks', IEEE Transactions on Communications vol COM-25, no 1, January 1977.

Gerla M, Monteiro J.A.S. & Pazos R.

'Topology Design and Bandwidth Allocation in ATM Nets', IEEE selected areas in Communications vol 7 no 8 Oct 1989.

Gitlin R.D. & Kaufield J.C.

'DataNetworking: An Issue Overview', AT&T Technical Journal, November 1988

Grout V.M, Sanders P.W. & Stockel C.T.

'Large Scale Telecommunication Network Optimisation' , Proc Applications Conference on the Simulation and Optimisation of Large Systems, Institute of Mathematics, University of Reading, UK, Sept 1986.

Heywood P.

'The Coming Transatlantic Price Wars', Data Communications, Jan 1991.

Heywood P.

'LAN-to-LAN Links in Europe: The Price Isnt Right - Yet', Data Communications, Apr 1991.

Jain R. & Routhier S.A.

'Packet Trains - Measurements and a New Model for Computer Network Traffic', IEEE Journal on Selected Areas in Communications col SAC-4 No 6 Sept 1986.

Johnson, Lenstra & Kan

'The Complexity of the Network Design Problem', Networks, vol 8, 1978.

Johnson J.T.

'Sonet for Private Netoworks', Data Communications International, Jan 1991.

Kleinrock L.

'Analytic and Simulation Methods in Computer Network Design', Spring Joint Computer Conference, 1970.

Kleinrock L.

'Queuing Systems Volume 1: Theory', Wiley & Sons, 1975

Kleinrock L. & Kamoun F

'Optimal Clustering Structures for Hierarchical Topological Design of Large Computer Networks', Networks vol 10 1980.

Lane, J.E.

'Packet Switchstream, PSS a Users Guide', NCC Publications, 1987.

McGregor P.V. & Shen D.

'Network Design: An Algorithm for the Access Facility Location Problem', IEEE Trans Communications vol COM 25 No 1 Jan 1977.

McQuillan J.M.

'Cell Relay: High Speed Networks Take a Step Closer to the 21st Century', Data Communications International, Sep 1991.

Mirzaian A.

'Lagrangian Relaxation for the Star-Star Concentrator Location Problem: Approximation Algorithm and Bounds', Networks, vol 15, Wiley & Sons, 1985.

Monma C.L. & Sheng D.D.

'Backbone Network Design and Performance Analysis: A Methodology for Packet Switching

Networks.' IEEE Journal on Selected Area in Communications, vol SAC-4 No 6 Sept 1986.

Moret B.M.E. & Shapiro H.D.

'Algorithms from P to NP' Vol 1, 1991, Benjamin / Cummings Publishing Co.

Mulvey J.M.

'Testing of a Large-Scale Network Optimisation Program', Mathematical Programming 15, p291-314, North Holland Publishing Co., 1978.

Norman H. Van.

'LAN/WAN Optimization Techniques'

Ozekici S.

'Queuing Theory and Applications', Hemisphere Pub Corp, 1990.

Pierre S. & Hoang H.

'An Artificial Intelligence Approach to Improving Computer Communications Network Topologies', J Ops Res Soc vol 41, no 5, 1990.

Saksena V.R.

'Topological Analysis of Packet Networks', IEEE Journal on Selected Area in Communications, vol 7 no 8 Oct 1989.

Sauer C.H.

'Queuing Network Modelling in the Design of Computer Communication Systems', IEEE, ICC 1978.

Schlar S.K.

'Inside X.25; A Managers Guide', McGraw-Hill, 1990.

Schwartz M.

'Telecommunications Networks, Protocols, Modelling and Analysis', Addison-Wesley, 1987.

Sharma R.

'T1 Network Design and Planning Made Easier', Data Communications, Sept 1988.

Singhal K, McAndrew C.C, Nassif S.R. & Visvanathan V.

'The Center Design Optimisation System', AT&T Technical Journal, May/June 1989.

Stalling W.

'Data and Computer Communications', 3rd edition, Maxwell Macmillan, 1993.

Tannenbaum, A.S.

'Computer Networks', 2nr Ed, Prentice Hall International, 1989.

Tarifica

'International Circuit Costs', Tarifica International, Intellidata Ltd, UK, 1993.

Tobagi F.A, Gerla M, Peebles R.W. & Manning E.G.

'Modelling and Measurement Techniques in Packet Communication Networks', Proc IEEE vol 66 No 11, Nov 1978.

Topkis, D.M.

'A k Shortest Path algorithm for Adaptive Routing in Communications Networks', IEEE Transactions on Communications, Vol 36, No 7 July 1988.

Whitt W.

'The Queuing Network Analyser', Bell System Technical Journal, vol 62, No 9, Nov 1983.

Whitt W.

'Performance of the Queuing Network Analyser', Bell System Technical Journal, vol 62, No 9, Nov 1983.

Yokoyama M, Miyake K & Nakajima I.

'Packet-Switched Network Planning System Architecture And Network Optimisation Algorithms.', Review of the Electrical Communications Laboratories, Nippon T&T Corp, vol 36 No 3 1988.

PUBLICATIONS

The following two papers have been submitted to Computer Communications for publication later in 1993.

Fast Heuristics for Practical TDM Network Design

G. P. Ray, AT&T, P. W. Sanders, University of Plymouth,
C. T. Stockel, University of Plymouth

Keywords : Network design, heuristic design, time division multiplexer, route search.

Abstract

Many large scale core data networks are now based on the new generation of intelligent time division multiplexers. These allow permanent virtual circuits between any network nodes and automatic rerouting should trunk failures occur. The design requirements for such networks are very different to those more usually associated with standard packet switched network arrangements. For a given traffic matrix the number of possible network topologies, link capacity selections, and circuit routings even between a small number of nodes is relatively large. When taking into account the additional requirements of optimizing available trunk capacity and ensuring resilience to failure by specifying link disjoint primary and backup paths the design complexity becomes very large. It is also necessary to use commercial trunk costs to ensure realistic design results.

The following paper describes the modules of a heuristic method that has been successfully used on a standard personal computer to provide minimum cost optimum designs operating under the above network characteristics. The heuristics incorporate a speed versus accuracy trade-off factor so that rapid approximate designs may be examined for differing traffic conditions.

Introduction

Many large scale networks are now provided using intelligent bandwidth management systems based on time division multiplexing (TDM). The intelligence of these systems is based on the operation of control software within each node whereby minimum 'cost' paths are selected, and the rerouting of channels is performed automatically in the event of a trunk failure. This may be achieved only if sufficient capacity is available on alternate paths, and it is usual for priority systems to be implemented to determine which channels are lost if insufficient capacity is available. These systems are different from the dynamic capacity sharing of packet switched systems in that they use dedicated capacity for each channel. In a packet based system it is the packet transit time, across the network, that degrades as capacity becomes limited following trunk failures. However, in TDM systems some calls may be lost while others maintain full service with the same delay when available capacity is reduced.

This paper looks at the design of large scale TDM networks, using a method developed for a standard PC yet still able to cope with designs of 20 or more nodes. This seems to compare favourably with other network design system [Agarwal, 1988]. in this design methodology traffic requirements are measured in terms of channels, typically 64kbps, for the sake of simplicity though it is possible to use any capacity units. The work has been applied to the design of European digital networks with trunk speeds up to 2Mbps.

In designing any computer network there are a number of trade-offs to be made in order to meet the objective of minimum overall cost. In selecting the optimal set of trunks from the myriad of possible choices, a rapid means is required for identifying those which are likely to be optimal. Merely selecting minimum cost trunks until a 'sufficiently' connected topology is reached or all traffic demands can be routed cannot be considered a successful design method. The definition of 'sufficient connectivity' is usually based upon a reliability factor required of the network. Typically, in order to meet a required service availability it is necessary to provide a minimum of two trunks connected to each node. The work described in this article assumes a minimum of two link disjoint paths are needed for all virtual circuits.

It has been recognised that the network design problem is difficult, Johnson, Lenstra & Rinooy Kan [1978] showed that it belongs to the NP-hard class of problems, the time to solution for such problems is not a polynomial function of the problem size. The solution time increases exponentially with increasing problem size, number of nodes, number of special requirements etc. For this reason heuristic techniques are investigated for a practical solution.

Many attempts have been made to design both packet switched and TDM networks using linear programming (LP) methods to create initial feasible topologies followed by perturbation methods to reduce the levels of redundant capacity [Frank & Chou, 1972; Gerla & Kleinrock, 1977]. The function of the linear program is to create an initial feasible topology of low cost. It is the quality of the perturbation schemes that one depends upon to drive the topology cost down to an optimal solution. The problem is that so many network topologies can be created with an initial 'low cost' and any attempt to drive one in particular to an optimum may find a 'local minima' in the sample space. The perturbation schemes must then be capable of generating a perturbation sufficiently large to escape any local minima without losing any 'optimal components' of the original design.

An alternative system to linear programming for finding initial feasible topologies is the Threaded Search, which describes a repeated application of a search for the next best route to allocate as the initial design is produced. The problem's difficulty then becomes one of determining what constitutes the basis for selecting the best route and dealing with the sheer number of possibilities. Even though the sample space for the general network design is very large, it is possible to develop a number of heuristic methods to 'weed out' the topology features that are, from experience and analysis, known to be poor. Such a system, being based on step by step trunk selection decisions, will suffer from a 'lack of foresight', making short term opportunist selections without regard for the global optimum solution. Using a method that makes the primary selection of routes on minimum cost (a greedy algorithm) does not lead to an ultimate minimum cost solution [Moret & Shapiro, 1991]. The use of lookahead techniques, examining the possibilities of all outcomes and their consequences, for a number of selections in advance, is well established. However, the Network Design Problem suffers from a great complexity and, though desirable, in aiming for fast heuristic methods, lookahead cannot be performed by brute effort directly.

The Threaded Search Structure

Feasible topologies may be produced by the systematic search and allocation of all paths, by evaluating their apparent cost, and then selecting the cheapest. The advantage of this method is that capacity is only sought for the paths having load requirements and the traffic is allocated by successively determining the cheapest load to carry. The 'thread' of the design process is therefore the order in which the selections are made, it is continuous for the initial topology design phase since no allocated path is evaluated for a second time until all traffic demands have been assigned a route and the required link capacity. In order to meet the requirement for resilience, to link failure on the primary route, it is necessary to include the allocation of capacity to a link disjoint backup route in the threaded search. In making the repeated selection of the next cheapest route it is possible to include the search for backup routes and allocate the cheapest whether found to be a primary or backup route. Since this is essentially still a 'greedy method' it is necessary to apply a heuristic technique to counter the lack of any lookahead. In order to determine how this might be achieved the route selection process is examined in more detail.

Selecting Routes

In order to see how a topology is created by a threaded search for best routes it is necessary to examine the way in which the cost function influences their selection.

The unit channel cost c between nodes i and j is given by;

$$c_{ij} = l_{ij} \, [\, k_0(\Omega_{ij}) + d_{ij} \, \Omega_{ij} \, f(\Omega_{ij}) \,] \; ;$$

$$where \; \Omega_{ij} = (\alpha_{ij} + \beta_{ij}) \, .$$

- d_{ij} is the distance from node i to node j;
- l_{ij} is a [0,1] matrix representing the presence of link i-j;
- Ω_{ij} is the allocated traffic variable representing the sum of primary and backup channels allocated;
- α_{ij} represents the primary traffic on link ij;
- β_{ij} represents the backup traffic on link ij;
- k_0 is the fixed link setup cost for the channel capacity Ω_{ij};
- $f(\Omega_{ij})$ is the cost rate of change function, it determines the change in per channel cost with increasing channel allocation.

To define the cost of a route, firstly, let a route vector R be defined as a series of L nodes between which the links form a single path such that

$$R = \{ \, r_1, \; r_2, \; r_3, \; . \; . \; . \; r_L \, \} \, .$$

Where links in the path denoted as r_x and x denotes the x^{th} link from node [x] to [x+1].

The absolute cost of a route is the sum of all included link costs, from r_1 to r_L. This assumes that no other connections are currently made in the network. Let the i^{th} link of the route be components that represent the link pair of the route,

$$c(R) = \sum_{i=1}^{L-1} [\, k_0\,(\Omega_{r_i r_{i+1}}) + d_{r_i r_{i+1}}\, \Omega_{r_i r_{i+1}} \,.\, f\,(\Omega_{r_i r_{i+1}})\,]\ .$$

The route cost is therefore dependent upon not only the total length but also the allocated capacity of each link making up the route. The allocated capacity will consist of two parts, that already allocated prior to the new route and that of the route itself which is a constant at the time of comparison between alternative routes for a chosen traffic demand.

In Europe the international link tariffs show a predominant drop in unit channel costs with increasing capacity. This leads to a downward force on costs as the allocated capacity of a link Ω, increases. Allocating more capacity to links, within limits, can therefore reduce overall channel costs at future stages of the design process.

However, the selection of minimum cost routes at each stage is desirable so long as they contribute to overall optimality. Where a small number of routes fall within a narrow cost bound the cost differential may be insufficient to justify selecting merely the very lowest cost without using a lookahead method. Under these circumstances an additional, or secondary criteria for selection may be employed. Returning to the route cost equations above, it can be seen that the second criteria for route selection may be used to lower unit channel costs by increasing the traffic allocated to links. Given a number of similarly priced routes the one that increases the traffic allocated across the network is therefore the one with the most links in its route.

The route selection process therefore consists of a search for all feasible routes and a repeated comparison that tests for those within a cost bound. This bound is referred to as the *'cost desensitivity'* because of its effect in reducing the route selection's sensitivity to minimum cost routes. The cost comparison is therefore based on the following test;

$$\left| \sum_{i=1}^{L_1} (\, K_0^{L_1} + d^i\, \Omega^i\, f\,(\,\Omega^i\,)\,) \pm \sum_{j=1}^{L_2} (\, K_0^{L_2} + d^j\, \Omega^j\, f\,(\,\Omega^j\,)\,) \right| \le \Delta$$

where Δ is the desensitivity factor.

The setting of this desensitivity factor has been found to be dependent upon the particular traffic and trunk cost conditions under investigation, but the requirement that the optimal setting be found proves very useful in network design projects. The desensitivity parameter gives the network designer a 'tunable' control with which to generate differing network topologies without needing to adjust the input data, i.e. traffic matrix and link tariffs.

The desensitivity factor controls the way in which additional capacity is encouraged in the topology to reduce the channel allocation. However, from the perspective of overall topology cost, there is a balance to be found towards the optimal level of additional capacity to be allocated.

The total network cost at any given stage of the optimisation, step n, can be written by replacing all static variables with the dynamic (n) time indices;

where the $l_{ij}(n)$ is a 0,1 matrix to represent whether a link from node i to node j is in place at stage n of the design process.

$$\psi(n) = \sum_{i=2}^{N} \sum_{j=1}^{i-1} l_{ij}(n) \,.\, [\, k_0(\Omega_{ij}) + d_{ij}\Omega_{ij}(n)\; f(\Omega_{ij}(n))\,]\;;$$

$$\text{where} \quad \Omega_{ij}(n) = [\, A_{ij}(n) + B_{ij}(n)\,]\;.$$

$$\text{Let} \quad K_0 = \sum_{i=2}^{N} \sum_{j=1}^{i-1} l_{ij}\, k_0(\Omega_{ij})\;,$$

hence

$$\psi(n) = K_0 + \sum_{i=2}^{N} \sum_{j=1}^{i-1} d_{ij}\Omega_{ij}(n)\; f(\Omega_{ij}(n))\;,$$

The diagram below indicates how, if the cost desensitivity can encourage the allocation of extra capacity for a controlled and small cost penalty, the overall topology cost benefits due to the lowered unit channel costs. This effect is only of positive where the cost penalty is small, if the desensitivity to cost is made too large then excessive capacity can be incurred that contributes to an increasing overall topology cost without lowering unit channel costs sufficiently to ensure a net cost improvement.

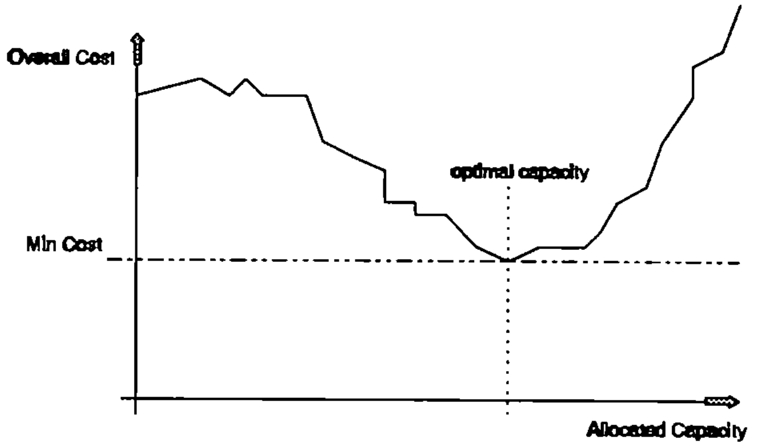


fig 1 Increasing total capacity allocation first improves, then degrades overall topology cost

The manner in which this deliberate increase in capacity can influence the topology is illustrated in figure 2 below.

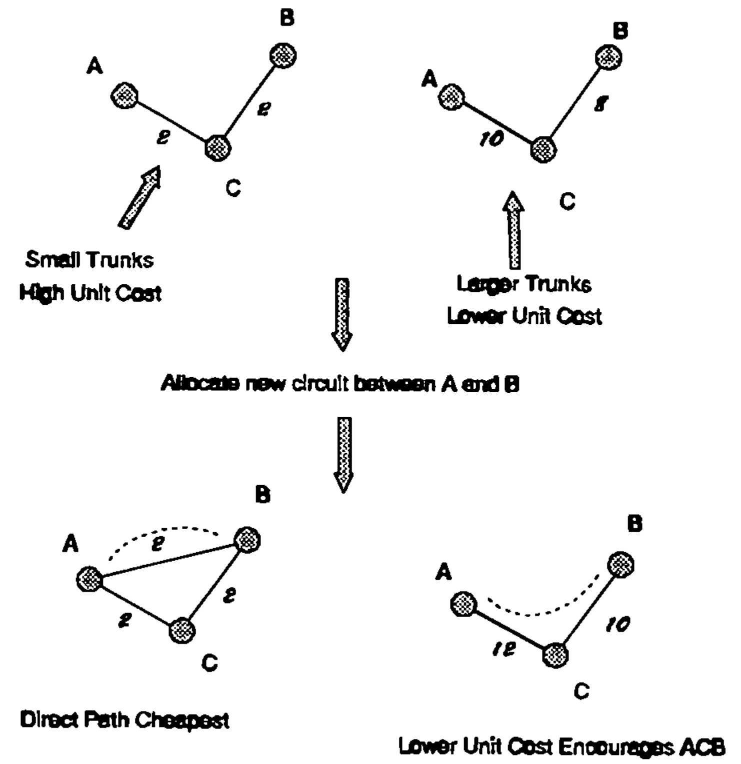


fig 2 Trunks of greater capacity have lower channel costs and can favour paths of greater length

To simplify the analysis consider the links AB to have zero capacity allocated initially. If the capacity allocation on links AC and BC is greater than this then assume that their unit channel cost is less than that of AB, per unit distance. The increase in distance that the path ACB may incur is dependent upon this difference in unit channel cost. Define the change in link i-j cost as $\delta_{ij}(D)$ where D is the change in allocated channels.
Thus

$$\delta_{ij}(D) = k_0(\Omega_{ij} + D) + d_{ij}(\Omega_{ij}+D)\; f(\Omega_{ij}+D) - [\; k_0(\Omega_{ij}) + d_{ij}\;\Omega_{ij}$$

In order that the bound on the excess distance via ACB be found, the cost of routes AB and ACB are equated, i.e. when the cost is the same, the limit of the boundary is defined. An expression can therefore be written where the boundary is defined as the point at which the cost increment for routes allocated A-B equals that for A-C-B:

$$\delta_{AB}(D) = \delta_{AC}(D) + \delta_{CB}(D) \quad ;$$

which may be expanded to

$$k_0(\Omega_{AB}+D)-k_0(\Omega_{AB}) + d_{AB}\ [\ (\Omega_{AB}+D)\ f(\Omega_{AB}+D) - \Omega_{AB}\ f(\Omega_{AB})\ ] =$$

$$k_0(\Omega_{AC}+D)-k_0(\Omega_{AC}) + d_{AC}\ [\ (\Omega_{AC}+D)\ f(\Omega_{AC}+D) - \Omega_{AC}\ f(\Omega_{AC})\ ] +$$

$$k_0(\Omega_{CB}+D)-k_0(\Omega_{CB}) + d_{CB}\ [\ (\Omega_{CB}+D)\ f(\Omega_{CB}+D) - \Omega_{CB}\ f(\Omega_{CB})\ ]\ .$$

Rearrangement produces an equation of the form:

$$K = d_{ij}\ v\ f(v) + d_{jk}\ w\ f(w)\ .$$

The right hand side now gives rise to a series of contours contingent upon d_{ij} and d_{jk} and the v and w values.

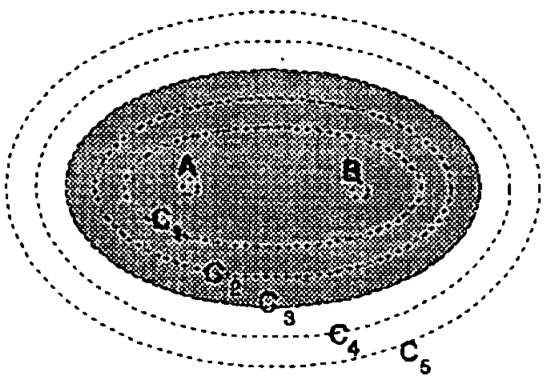


fig 3 Ellipses of constant inter-nodal distance

Figure 3 shows the range of possible positions node j may take, for each capacity C_1 to C_5, to be within the region to consider an alternate route. The boundary therefore represents the cost contour of each capacity. It is possible to examine any given network and evaluate the possible cost trade-offs of all possible routes, and hence the potential saving, by selecting routes with the greatest number of links to gain overall cost savings from increased capacity and attendant reductions in unit channel cost.

This may be extended to the variable case where the links AC, CB and AB may take on different capacity values when the boundary of the ellipse will change. Between any pair of nodes with a given link capacity it is possible to calculate the maximum extra distance beyond the direct path length that the route may take. This takes into account the unit channel cost reductions gained by sharing other nearby links that are available, and gives rise to the regions within which intermediate nodes may be used for routing. For a simple single intermediate node, assuming equal channel costs for the two link alternative route, the following applies:

$$X \cdot c_1 \leq Y \cdot c_2\ ;$$

where c_1 = unit channel cost of the direct path;
c_2 = unit channel cost of links between intermediate nodes;
X = distance between nodes A and C;
Y = total distance between A,B and C.

In order to greatly reduce the number of routes to be searched it is possible to limit the nodes examined for possible intermediate hops. If the route-finder complexity is to be reduced it is necessary to find a fast method for determining a reduced set of N. The net gain in speed must not be lost in the 'speed up ' algorithm itself.

It has been noted that the most likely regions in which routes are formed between two nodes mark an approximate ellipse around the end nodes. It is possible to very quickly determine the 'ellipse nodes' to be considered by the route-finder by using the following test. For end nodes A and B separated by a distance d_{AB} consider a candidate node X, distance D_{AX} and D_{XB} from nodes A and B respectively. The ellipse is defined as a factor *d* of the AB node separation, for instance Node X is within an ellipse $d.D_{AB}$ if:

$$d.D_{AB} >= D_{AX} + D_{BX}$$

In practical design situations for international networks, trunk tariff structures bear little resemblance to physical distance and it is possible to replace all distance variables with those of corresponding cost. The 'elliptic bound' when used in route searches has been found to rapidly identify good candidate intermediate nodes and has an increasing benefit as the number of network nodes increases. The complexity of the route searching algorithm, when examining all network nodes, N, is a function of the number of permutations of M nodes from N, where M is the maximum number of intermediate nodes. If however the 'elliptic bounding' is used to select E candidate nodes the number of permutations is a function of the number of combinations of M from E. where $M \leq E \leq N$.

Perturbation Schemes

Due to the threaded route selection process and discrete trunk capacity availability it is possible that network capacity is allocated in a less than optimal manner and any excess may be identified using perturbation schemes. Perturbation schemes aim to select high cost (or low efficiency) trunks and replace them with cheaper (or more efficient) alternatives. The means for identifying single or groups of trunks for deletion is usually made on the basis of trunk utilization. Some approaches such as the Branch Exchange algorithm or the Cut Saturation [Gerla & Kleinrock, 1977] look for minimum efficiency trunks to delete and the minimum cost alternatives to restore the traffic loss.

Two perturbation schemes have been found to be particularly effective for use with the threaded design method. The first is based on identifying any trunk with less than 75% utilization. A list is built up of such trunks and they are deleted in turn, the displaced loads are then rerouted according to the minimum cost path available. It is often found that the original trunk is restored, but savings are consistently found for a large number of topologies. The second perturbation scheme identifies any trunk to which a single path is allocated. The link is deleted and the minimum cost alternative path is sought. Very often it can be found that sufficient spare capacity exists within the topology to accommodate the routes displaced by the trunk deletion.

The problem with each of these schemes is that, where more than one candidate trunk is found for deletion there is the problem of selecting which order to perform the delete and add. It is necessary to repeatedly run the perturbation systems until no further improvement is found and the ordering problem can mean this extends the run time.

Results

The speed advantage of the elliptic bound is illustrated below. The tests were performed on a low powered 386DX IBM-PC. The following variables are used:

P =	maximum number of hops allowed in a primary path;
B =	maximum number of hops allowed in a backup path;
E_{max} =	elliptic bound, number of candidate nodes for route search in addition to P or B for primary or backup routes respectively.

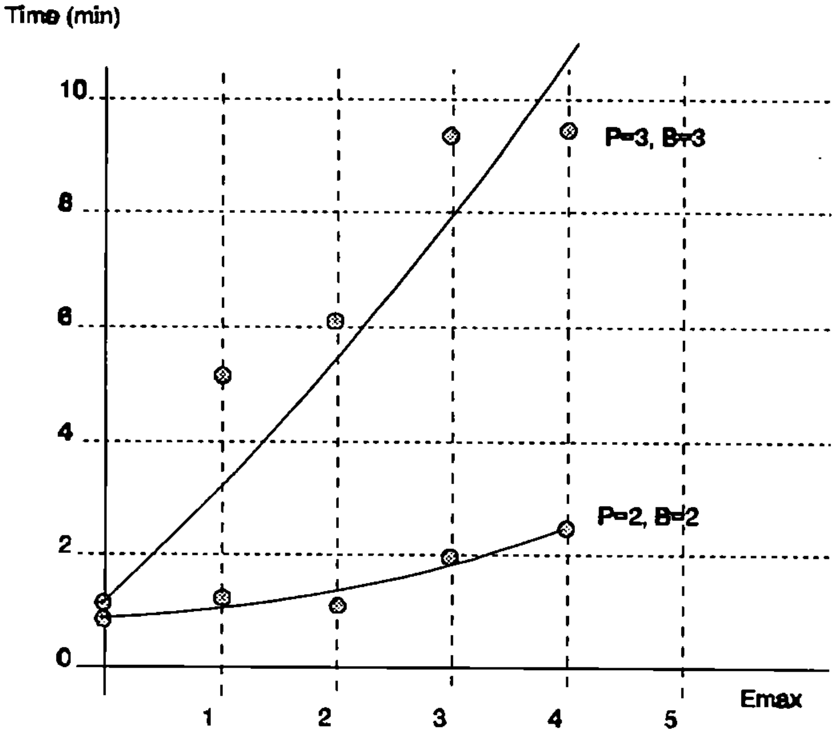


fig 4 Graph to show the design times for 8 node topology

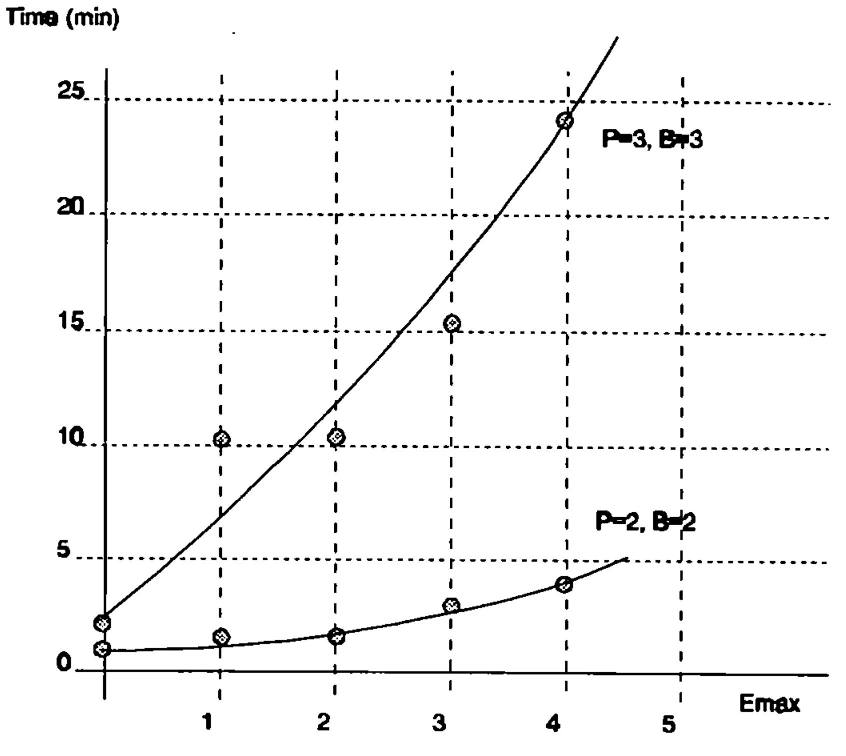


fig 5 Graph to show the design times for 9 node topology

The routes are searched in order, starting with 1 hop paths increasing up to P_{max} maximum primary hops and B_{max} maximum backup hops. The elliptic bound is described by the value E_{max} which represents the maximum number of nodes between which to search for routes, in addition to those allowed for by the P_{max} and B_{max} parameters. The number of nodes used in the route search is therefore a maximum of P_{max} + E_{max}:

e.g. if P_{max} = 3 and E_{max} = 2, then the elliptic bound is extended to encompass the 5 nearest nodes to the end points falling within an ellipse.

In terms of computational complexity this therefore implies that the work required to search for maximum 3 hops routes with E_{max} = 2 is equivalent to searching for 5 hop routes with E_{max} = 0.

The cost improvements as a result of expanding the maximum number of hops per route is illustrated in the tables below. The design times indicated reflect the total time taken to create the initial topology and then perform perturbations until no further topology cost improvement was obtained.

Nodes	Hops	E_{max}	Cost	Design Time Secs
8	2	0	116,114	41
		1	100,219	94
		2	108,114	84
		3	108,114	109
		4	108,114	132
	3	0	100,219	93
		1	102,577	339
		2	101,382	370
		3	91,212	565
		4	91,212	565

Table 1 N = 8, Cost and Design Time

Nodes	Hops	E_{max}	Cost	Design Time Secs
9	2	0	139,834	63
		1	132,294	146
		2	134,663	131
		3	134,663	184
		4	134,663	233
	3	0	132,294	146
		1	129,570	612
		2	127,388	640
		3	116,923	1012
		4	116,923	1437

Table 2 N = 9, Cost and Design Time

The effects of the desensitivity factor on network cost are shown in table 3 below. While there is no direct relationship between desensitivity and minimum network topology cost, there is a definite range of values for each topology that yields improved designs. Additional columns are shown to indicate the cost of the capacity allocated but unused, the percentage cost utilization and percentage capacity utilization. The last column shows that ratio of capacity allocated to the primary path against that for the backup path.

Topology	Desens (%)	Cost (£)	Spare Value (£)	Cost Util (%)	Capacity Util (%)	P/B
T-1	0	151,476	30,969	79.6	76.6	1.57
T-2	10	145,417	33,834	76.7	72.6	1.11
T-3	20	151,083	22,270	85.3	83.0	1.11
T-4	30	140,556	22,153	84.2	81.1	0.90
T-5	40	144,553	18,166	87.4	83.9	0.74
T-6	0	146,291	21,116	85.6	83.8	1.38
T-7	10	145,873	30,347	79.2	75.2	1.31
T-8	20	141,002	23,848	83.1	80.0	1.37
T-9	30	139,284	20,480	85.3	80.2	1.20
T-10	40	141,711	22,428	84.2	81.0	1.08

Table 3 A range of 14 node network solutions

Conclusion

The core network design problem for TDM systems has been shown to be tackled using an effective new system. The consistency of the design quality gives an indication that the design methods are soundly based. Good network designs have been generated, to a quality suitable for use by a major international network provider. The new method offers an alternative solution to a problem previously only tackled using systems based upon linear programming techniques and requiring considerable mainframe processing power. Two new techniques allow similar sized problems to be tackled on medium power personal computers. The first new technique, called 'cost desensitivity', improves the quality of designs using greedy algorithms by sacrificing short term gains for a global topology objective. The second new system used in route selection is called 'elliptic bounding'. This greatly reduces the execution time of searches for feasible paths by forming a logical limit to the number of nodes between which routes are tested. Using the simple equation of the ellipse the test itself adds an insignificant overhead to the overall calculation. The speed improvements of the elliptic bound are related to the size of the bound, the greater the speed the less likely are the routes to be optimal. However, network designs can be constructed rapidly using fast search methods and cost estimates given to within typically 10% of a complete solution's final cost using a slower, more comprehensive search.

The designs can be shown to satisfy all given constraints and design costs have been exhibited within limits acceptable to the end user. The results are generated on a medium power IBM-PC within timescales (typically 10 minutes to 5 hours) acceptable to the designer who is given full control over network parameters and the opportunity for manual adjustment. Cost optimality is not guaranteed but, in order that differing network designs may be compared qualitatively, detailed figures are given showing the breakdown of both financial and capacity utilisation for all network loads. become a max acceptable transit delay.

It would be possible to design packet switched networks using the same techniques, though instead of calculating minimum cost paths it may be preferable to evaluate routes on the basis of both minimum delay and cost. Trunk utilization would then be a non-linear function of load and the 'full' limit would therefore be based on delays exceeding defined thresholds rather than trunks becoming full in the case of TDM systems.

Acknowledgement

The authors wish to thank John Allen of AT&T BCS-E for his advice and suggestions in the preparation of this manuscript.

References

1. Agarwal Y.K.

'An Algorithm for Designing Survivable Networks', AT&T Technical Journal May/June 1989.

2. Johnson, Lenstra & Rinooy Kan,

'The Complexity of the Network Design Problem', Networks, vol 8 1978

3. Frank H. & Chou W.

'Topological Optimisation of Computer Networks', Proc IEEE vol 60 p1385-1397 Nov 1972.

4. Gerla M & Kleinrock L.

'On the Topological Design of Distributed Computer Networks', IEEE Transactions on Communications vol COM-25, no 1, January 1977.

5. Moret B.M.E. & Shapiro H.D.

'Algorithms from P to NP' Vol 1, Benjamin / Cummings Publishing Co. 1991.

A New Technique for the Representation of Access Network Performance

G. P. Ray, AT&T, P. W. Sanders, University of Plymouth,
C. T. Stockel, University of Plymouth

Keywords : Network Performance, Multiplexer Delay, Performance Estimation

Abstract

Many large data communication systems can be conveniently divided into core and customer access network arrangements. The design of the latter, with statistical multiplexers and leased line or dial-up access is often neglected in terms of optimum design. This is simply because the attachment of users is affected in an ad-hoc manner over a period of time and is highly susceptible to physical changes and variability in traffic demand. More attention is now being given to this relatively expensive resource, especially as the quality of service in terms of delays and blocking are an increasingly important part of service level agreements between the value added network supplier and customer. This paper presents a new technique, suitable for a wide range of queuing systems, that aids the optimal design of star type networks where the quality of service constraints must be taken into consideration.

Introduction

In the field of wide area data networks (WANs) there is usually a clear distinction between the access portion of the network, often an arrangement of multiplexers through which customers gain connection, and the core network that supports the intercommunication of users' data over a large distance.

In assessing the level of performance expected of a network during the design phase, the core network delays are often assumed to be fixed and the delays specific to the access multiplexer and node switches are estimated as a function of the offered load. Controlling the network delay is an important part of ensuring a guaranteed level of service and forms an important part of customer service level agreements. The performance analysis is effectively broken down into two separate parts where the complexities of the combination of various traffic sources at each access point are divorced from the interconnection problems of the core network design. Access points are considered in isolation in order that the delays related to the multiplexer queuing can be assessed in relation to the connected traffic loads [Al-Chalabi M. & Liss W.J, 1988].

The core network design then assumes a number of simple traffic requirements based on the aggregation of all traffic demands at each access node [Pierre & Hoang, 1990]. This distinction is simple to deal with in many WANs since the access networks are usually based on packet technology and the core networks commonly now use time division multiplexing to share large bandwidth trunks between a number of different data services. These services might typically include simple point to point connection, frame relay, SNA and X.25.

The following analysis and developed techniques have been based upon the characteristics and operational experience gained with the AT&T Business Communication Services - Europe (BCS-E) Infotrac access network.

The Network

The Infotrac network provides low speed asynchronous data communications facility with dial-up access at local telephone call charge rates for speeds up to 1200 baud. Direct connection, using leased lines, to the largest customers with speeds up to 2400 baud is now becoming a standard.

The services offered across the Infotrac network are predominantly viewdata based and typical examples are the mortgage, insurance and holiday quotation systems employed by many high street brokers.

The viewdata service presents a mixture of coarse graphics and textual information. The data is generated by central host mainframes running the various holiday, insurance, etc quotation systems. The user will usually type in a small number of keystrokes to select various menu options and enter a service specific response for such details as customer name and any product information. The traffic is therefore typified by large blocks of data, up to 300 characters at a time, from the host to the user, and by a small number of characters from the user to the host in response.

The equipment supporting the Infotrac service is supplied by Digital Communication Associates (DCA) now part of the Racal Electronics Group. The network configuration is based on local statistical multiplexers (muxs) in each main UK town or city to which is connected a bank of modems for dial-up customers and leased lines for 'direct connect' customers. These essentially form the access layer, which is then further concentrated into regional switching muxs between which the region to region core network is constructed.

The mux to mux communication is implemented using a proprietary DCA protocol based on the Digital Equipment Corp (DEC) Digital Data Communications Message Protocol (DDCMP). This operates using frames and sub-frames, called plexs. A plex is used to carry the characters to and from each terminal connected to the mux, and frames are used to carry the assembled plexs from all connected terminals at that access point. A frame is transmitted when full, or if not full after a predetermined frame forwarding time (FFT). The FFT is a parameter that is tuned by the network provider to ensure optimum operational conditions within the network. It requires a trade-off between a time sufficiently long to ensure full packets and minimise protocol overhead and one sufficiently short to ensure adequate response times to the user.

The main users' perception of network performance is based on the response time to provide character echo to keys pressed on the user terminal. The consistency of network delay is also an important factor since the drawing of screens of data on the user terminal is the result of the transmission of a large number of characters. Approximately constant delays make the drawing of text screens appear smooth, but when erratic, the screen appears to be drawn jerkily and the user may find this disconcerting. User trials show that a consistent echo delay is desirable, even if a slightly larger delay is incurred, in preference to a low mean delay with large fluctuations. Thus, if delay performance is to be predicted it is as important to determine the delay profile as it is to identify the mean delay.

Since Infotrac is based on a mainframe host service the character echo is generated by the host machine (host echo) rather than the access network multiplexer or switch. In order to identify the major influences upon network delay performance it is necessary to identify the contributory elements to the overall character echo.

Within the process that creates the character echo, the dominant delay is that due to the access multiplexer/switch on the return path of the character from host to user. The only traffic within the access network on the user to host path is due to keypresses from users, and these occur at a low rate. In contrast, the host to user echo characters are interspersed with the large screens of data that may be sent to other users on the same portion of the access network.

One of the problems with designing this type of access network is that there are customers from a variety of different service classes attached and there is no single consistent traffic profile for the inputs. However, they can be broadly split into a number of market sectors, within which there is a strong correlation of traffic profile, mainly due to the type of network based application in operation. For example the financial market offers users a range of life insurance and endowment policies that are presented to customers in the form of lists, from which a number may be picked and examined in more detail. The traffic profile of this type of session is consistent within the finance sector yet distinct from the leisure industry. A travel agent may scan through a range of holiday options on offer, discussing each in turn with the customer. A different yet consistent traffic profile for the leisure industry results.

Also, within the Infotrac network each market sector only uses a relatively small number of host mainframe applications. Consequently there is a high degree of commonality between the data sent to the customer, within each sector, in the form of viewdata screens of text and graphics.

Since a value added network provider usually offers network services for a number of market sectors there will be a mixture of customers from these market sectors on each access multiplexer. The access design problem faced by the network provider requires that the market sector of each customer be taken into account. It is possible to control the number and market sector of the customers connected to each multipexer by selecting the appropriate direct connections and dial-up access. The distribution of different access multiplexer phone numbers to different customers acts as a relatively simple and flexible way of controlling traffic. The design must take account of the different customer market sectors and finding the optimal mix on the multiplexers requires an understanding of how the various combinations of customer influence access delay performance.

The Design Technique

A new method for representing the access network performance is based on the combination of a small number of basic traffic models that represent each customer market sector. The models may be combined in the number and proportion of a proposed access multiplexer configuration and the effect of differing data rates and protocol setting examined. One of the most critical protocol parameters is the frame forwarding time, during which all incoming characters are buffered, and hence delayed, is included. From the combination of all the traffic profiles a three-dimensional graph is plotted showing delay versus traffic levels (arrival rate and frame/message length). A probability of each traffic level is then used to determine the overall likelihood of the multiplexer delay bettering a range of delay thresholds. From this a two-dimensional graph may be plotted showing the percentage probability of the delay being within all possible ranges. It has been shown [Grout & Sanders, 1989] that, in general, the Infotrac messages follow the standard Poisson arrivals and negative exponential message lengths to a very good approximation.

In order to find the total composite arrival rate for a number, N, independent arrival rate processes at a multiplexer direct addition of the rates can be used. The sum of Poisson distributions produces a new Poisson distribution whose mean is equal to the sum of the individual means [Ozekici, 1990].

The traffic profiles can be used in the summation process to quickly find the aggregate expected mean arrival rate. If a multiplexer had 10 Motor Industry users, 12 Leisure users and 5 Finance user then the total traffic arrival rate would be

$$\lambda_{Total} = 10.\lambda_M + 12.\lambda_L + 5.\lambda_F$$

where λ_M = Mean Motor user arrival rate
λ_L = Mean Leisure User arrival rate
λ_F = Mean Finance user arrival rate

Similarly the overall aggregate message length distribution for the users can be obtained from the weighted sum of the individual message length distributions. For the continuous distribution:

$$\frac{1}{\overline{\mu}} = \frac{1}{27}\left[\frac{10}{\mu_1} + \frac{12}{\mu_2} + \frac{51}{\mu_3}\right]$$

In representing these functions in software for the model a sampled distribution is necessary. This was calculated on a piecewise basis, to find the probability of each message length for every element, i, of the sampled distribution. This gives an approximation of the composite mean λ and $1/\mu$ for the combined inputs of a multiplexer.

It is now possible to build up a single message length and arrival rate model based on the composite mean λ and $1/\mu$ values for the combined inputs of the multiplexer. The figure below illustrates how the composite traffic model is built up by combining the basic traffic profiles for each market sector.

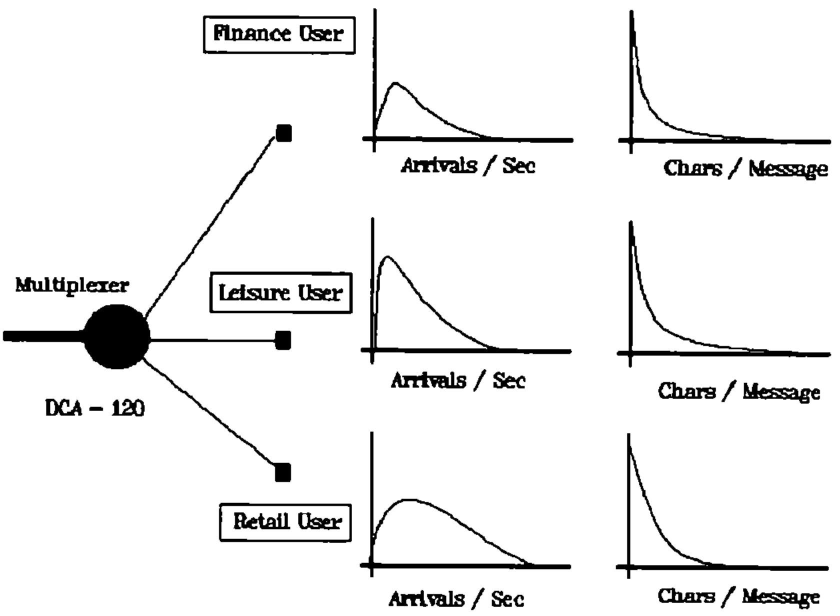


fig 2 Illustrating the combination of a number of user traffic profiles

The internodal protocol increases the volume of data on the DCA links by adding a management data overhead. The additional data added by the DDCMP protocol consists of one header per frame and a further one per plex. In order that the multiplexer delay figures represent the effects of the internodal protocol it is necessary to calculate the number of extra characters due to DDCMP at the time the total traffic entering the multiplexer is calculated.

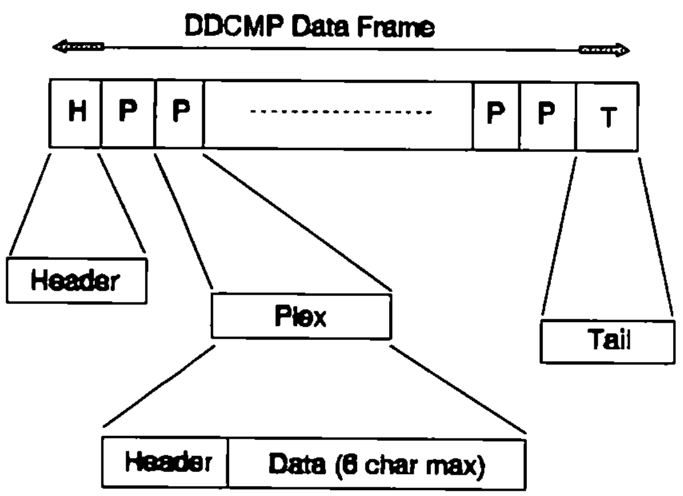


fig 3 The format of a DDCMP protocol frame

The total additional protocol characters added to the original user traffic profile is based on the actual values of frame (message) length and arrival rate over the range of values they take. Integer arithmetic is used to calculate how many data overhead characters would be produced for each possible arriving message length for all the users connected to the multiplexer. The basic traffic model is then adjusted to account for this increase in traffic.

Referring to figure 3 for the DDCMP protocol:

- h = Plex header length;
- P = Plex length;
- F = Frame envelope length;
- H = Frame Header Length.

Define a function to account for the additional plexs or frames that are only partially filled, relating it to the result of the 'modulo' (MOD) function:

K is a variable that can only take on the value 0 or 1;
K = 0, if the MOD function has a non-zero remainder;
K = 1, if the MOD function has a zero remainder.

In order to write an expression defining the total number of characters leaving the multiplexer output, the basic input arrival rates must be known and the number of additional characters created by the protocol must be calculated. To simplify the analysis, consider only a single input, with arrival rate λ and message length σ_c characters, or σ bits. Where $\sigma = 1/\mu$ and μ represents the message service rate in message/bit.

The total overhead characters per second due to the protocol will be due to the creation of full and partial plexs and the resultant frames.

i. Define N_c to be the total number of characters, on average, offered by the multiplexer inputs for insertion into a protocol envelope per second.
N_c = total number of data characters + number of characters due to plex headers, total data chars per second = arrival rate * message length = $\lambda.\sigma_c$:

$$N_C = \sigma_c.\lambda + \lambda.\ [\ (\sigma_c \text{ DIV } P) + K\ ].h$$

The characters due to the plex headers are caused by each arrival being assigned to its own plex. Therefore a minimum of one plex per arriving message is required. The number of plexs per message is defined by the message length divided (using the integer divide function DIV) by the plex size + one extra plex to carry the remainder of characters not exactly filling a plex (hence the K term), if one exists.
The overhead due to the plex headers = $\lambda.[(\sigma_c \text{ DIV } P) + K\ ].h$

ii. Define the total number of protocol frames per second to be N_e,
N_e = total number of characters (including plex headers) divided by the frame size, the K function means that the partially filled frame is accounted for.

$$N_e = [N_c \text{ DIV } (F\text{-}H)\ ] + K$$

iii. Let Δ be the total character overhead due to the protocol,

$$\Delta = \lambda.[(\sigma \text{ DIV } P) + K].h + Ne.H$$

In order to estimate the delay for the range of traffic arrival rates and message lengths from the sampled distributions an assumption is made. In essence, it is necessary to calculate the expected delay for each and every combination of λ and $1/\mu$ from the sampled distributions. This effectively means that for all possible input traffic levels the queuing delay is calculated.

Using the modified traffic model, the mean queuing delay, E[T], is given by the standard expression;

$$E[T] = \frac{1}{\mu . C - \lambda}$$

where the output channel transmission capacity is C bits/sec.

In order to do this it is necessary to assume that each value from the sampled distribution may be used independently. This requires that the samples themselves are considered to be the mean values of separate distributions. It is accepted that the use of the assumption may introduce some error but the form of this analysis and the results it produces are significantly more useful than a single mean value result.

The delay is therefore calculated on an elemental basis for the range of all λ and $1/\mu$. This will naturally result in a three dimensional surface with axis in delay, λ and $1/\mu$. The surface can be considered to be the locus of the operating point for the network traffic, i.e. as the network traffic load changes, the arrival rate and message lengths are continuously changing and the surface describes the expected delay for the instantaneous operating point. This is not an entirely accurate picture of the system, because at time t_n the locus does not take account of the queue's state at time $t_{(n-1)}$. Again, this may be considered to be an acceptable inaccuracy because the manner of the analysis and the range of the results provided are useful.

Explanation & Interpretation of results

A Network Analysis (NETAL) software package has been written which allows the storing of traffic arrival rate and message length profiles on disk and the user to create models of multiplexer inputs based on mixtures of network customers from differing market sectors. The multiplexer protocol and aggregated traffic profiles are used to estimate likely queuing delays within the network.

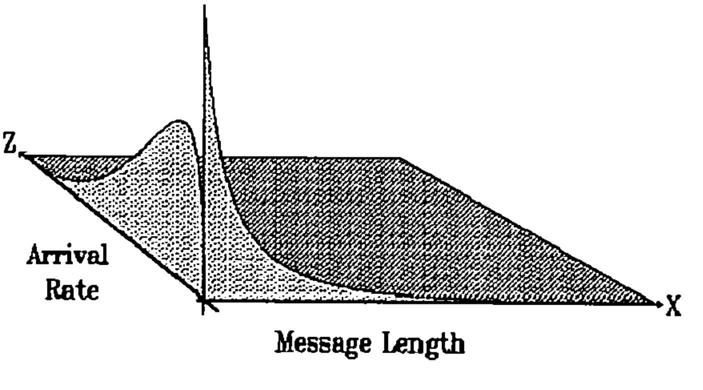


fig 4 Plotting arrival rate and message length

The NETAL package calculates and displays the expected delay figures as a three-dimensional (perspective) plot, the x axis represents the message length and the z axis the arrival rate. The y plane shows the expected delay. The graphs may be interpreted basically as follows:

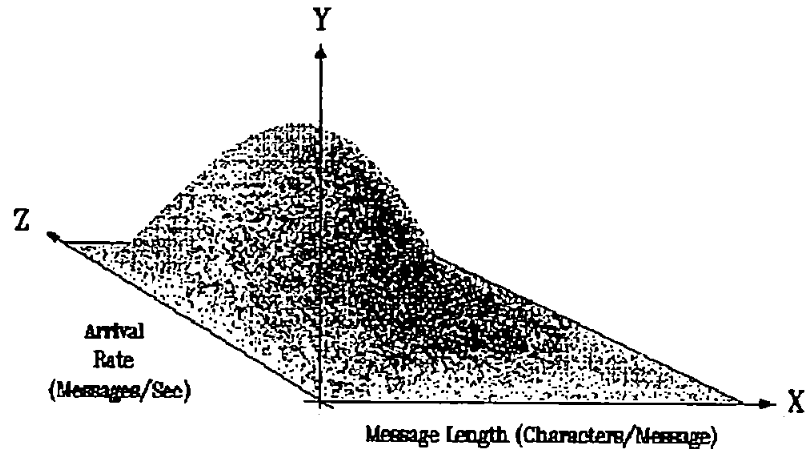


fig 5 Illustrating a smooth transition from one delay figure to another for varying traffic levels

If the delay surface appears smooth one may deduce that, as the traffic levels change, the operating point moves gradually from one delay value to another. The access network may therefore be said to be working 'smoothly', there are no rapid, large changes in delay.

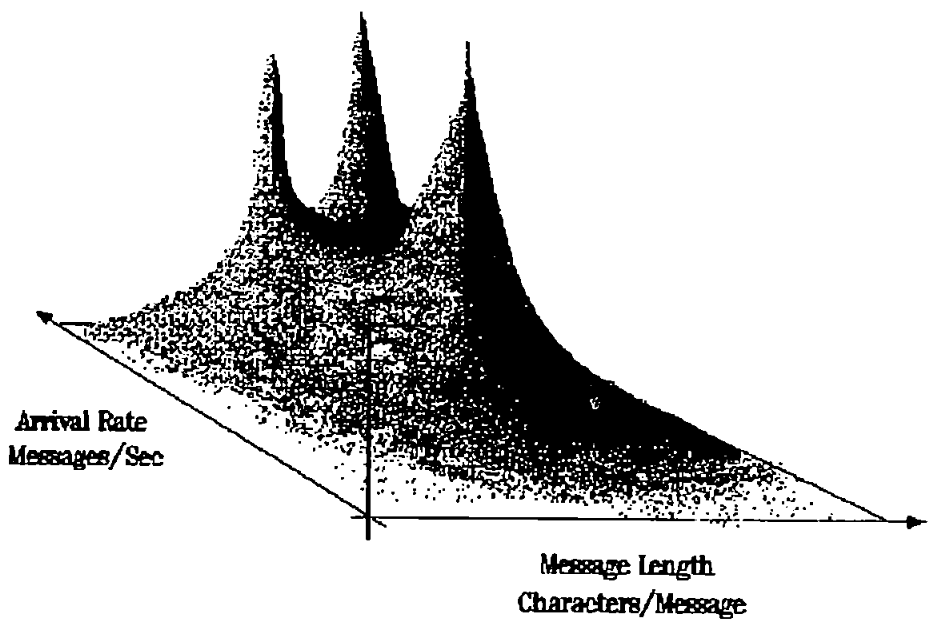


fig 6 Illustrating a rapid change in delay for small changes in traffic levels

A heavily peaked surface represents the rapid transition of time delay from one value to another as the traffic levels fluctuate by small amounts. The network would suffer large changes in delay and would appear congested to the user as some traffic takes substantially longer to return from the host than others. In practice the network designer requires a less complex representation of the delay performance for any given multiplexer configuration, with some means of estimating the percentage probability of the delay being less than a pre-defined threshold. For this the probability of each possible arrival rate and message length combination occurring are required.

Using the Joint Distribution Function

The view of the 'delay surface' must be tempered with the realisation that the delays represented by each point on the graph are not equiprobable. That is, the probability of each combination of message length and arrival rate will dictate the probability of the traffic operating point being at each point on the surface. This combined probability of each arrival rate and message length is defined by the Joint Distribution Function (JDF).

If the JDF of the total input traffic is known and the delay figures as a result of each traffic level are also known then it is possible to calculate the total probability of the delay being within specified ranges. Since the delay is defined on the y axis the surface may be intersected at any point in the xz plane and the area of the intersection will indicate all values of arrival rate and message length which lead to a delay above the threshold. These values of message length and arrival rate can then be tested against the JDF to reveal their probability of occurrence.

In order to visualise the problem, if the network designer wishes to know the probability of the delay exceeding a threshold then he may consider a horizontal slice being made through the delay graph at the level of that delay. The section of the graph above the cut represents the delays exceeding the threshold and that below it where the delay is better than the threshold.

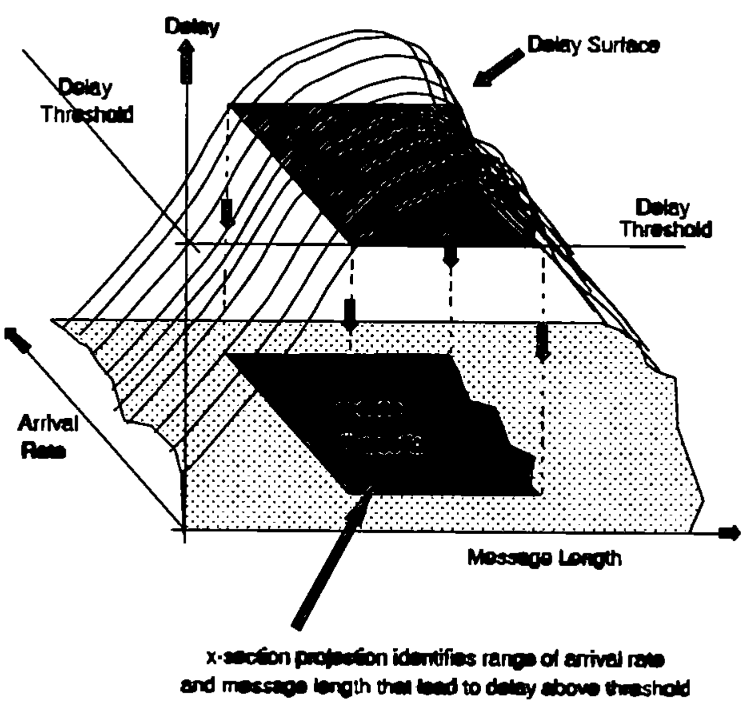


fig 7 Slicing the delay surface to find probability of threshold being exceeded.

The values of message length and arrival rate (MLAR) for which the delay threshold are exceeded is projected onto the zero delay axis. Mathematically this is performed with the Kroenecker function, which is used as a simple switching decision function taking on only two integer values either 0 or 1.

$$\Gamma\,(\lambda, 1/\mu) \quad = 0 \ \text{ for delay} > \text{threshold}$$
$$\qquad\qquad\qquad = 1 \ \text{ for delay} <= \text{threshold}$$

The salient feature of the JDF is that it describes the entire probability range of all (100%) occurrences of traffic loading, i.e. the volume under the surface must sum to unity. This is an important point to make because it is the central argument in estimating delay probabilities using this technique.

In order to calculate the percentage of the time that the time delay exceeds a certain value, the Kroenecker function is now applied to the JDF. The joint probability of the MLAR traffic values are summed for all '1' entries in the Γ function. If every delay was greater than the threshold then every value of Γ would be '1' and the probability would sum to unity, i.e. 100% probability. Consequently, for only the values of delay that exceed the threshold, where Γ = 1, the JDF values are summed to yield the total probability. It is also important to remember that both the message length and arrival rate probabilities tend to zero and the extremes of the area under the surface. Since the JDF is also the product of the two, the importance of the extremities is very small and from the results they may be safely neglected without leading to any significant distortion of the probability figures. For example, if the probability of message lengths greater than 1000 characters is of the order of 1:10,000 and the probability of their arrival rate being less than 0.1 per second is of the order of 1:100,000. Though this arrival rate and message length might lead to a delay of 14 seconds (assuming 7 bits/char and 1200bps line speed) the probability of its occurrence is $1{:}10^{-9}$.

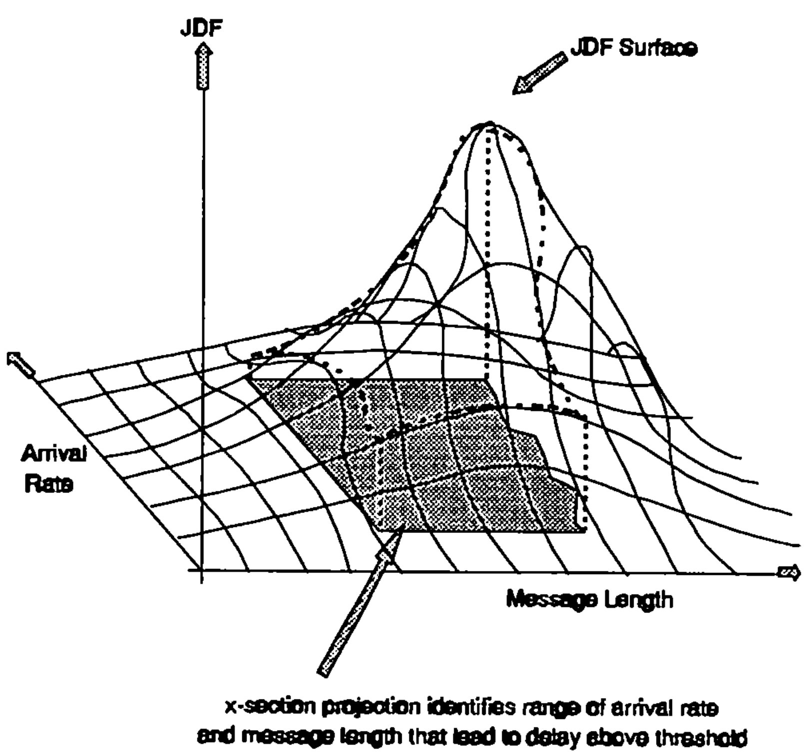


fig 8 Projecting the arrival rate & message lengths leading to delay above threshold onto the JDF

In summary the method is as follows:

i. On the delay surface, note all the combinations of arrival rate & message length that lead to a delay exceeding T milliseconds;

ii. sum all JDF probabilities for points where the corresponding delay exceeds the threshold T.
Define a binary function $\Gamma(\lambda, 1/\mu) = 1$ for $D(x,y) > T$,
0 otherwise;

iii. calculate the percentage probability of delay > T .
i.e. the total volume under JDF found in step ii;

iv. repeat from step i for a range of time delay thresholds, $T_1..T_{max}$.

Results

A two axis graph can now be plotted showing delay thresholds, T, against the probability of the traffic delay exceeding each threshold. The NETAL package has produced many such graphs and the two examples presented in the results section clearly show that the percentage based delay figures gives a good indication of the access network performance. Result 1 shows that, for approximately 80% of the time, the delay is less than 282mS yet for 95% of the time it is less than 300mS. i.e. the variability of the delay is not large. In contrast, result 2 shows a delay of better than 791mS for 79% of the time but a delay of better than 1230mS for about 95% of the time. The dramatic swing in delay within the same percentage bound would indicate a much more variable delay performance and shows that the worst case figures are likely to be unacceptable to the user.

Conclusion

The results of this work show that it is possible to view the way in which access multiplexer delay is influenced by the mixture of traffic from different customer market sectors and the protocol parameters. The graphical results allow the visualisation of the multiplexer performance and the simple graph of percentage probability of delay less than X mS is very useful to the network designer. In normal day to day network operations, when performance problems are identified, it would be possible to adjust either the customer 'mix' or the frame forwarding timer in order to bring delays within acceptable limits.

Further work is required to incorporate this method into the access multiplexer location problem. Using first pass methods such as the Add and Drop [Bahl & Tang 1974] or the general concentrator location methods of Narasimhan Pirkul [1989] it should be possible to ensure expected performance is within specified bounds for given customer locations and traffic profiles.

Acknowledgement

The authors wish to thank John Allen and Mike Fennell of AT&T BCS-E for their help and suggestions in the preparation of traffic profiles and market sector modelling .

References

1. Al-Chalabi M. & Liss W.J.

'Design of the Bank of America California Data Network', AT&T Technical Journal Nov/Dec 1988.

2. Bahl L.R. & Tang D.T.

'Optimisation of Concentrator Locations in Teleprocessing Networks', Symposium on Computer Communication Networks and Teletraffic, Polytechnic Institute of Brooklyn, April 1972.

3. Grout V.M. & Sanders P.W.

'Infotrac Traffic Profiles by Market Sector', Research Report, Dept Communication Engineering, Polytechnic South West, 1989.

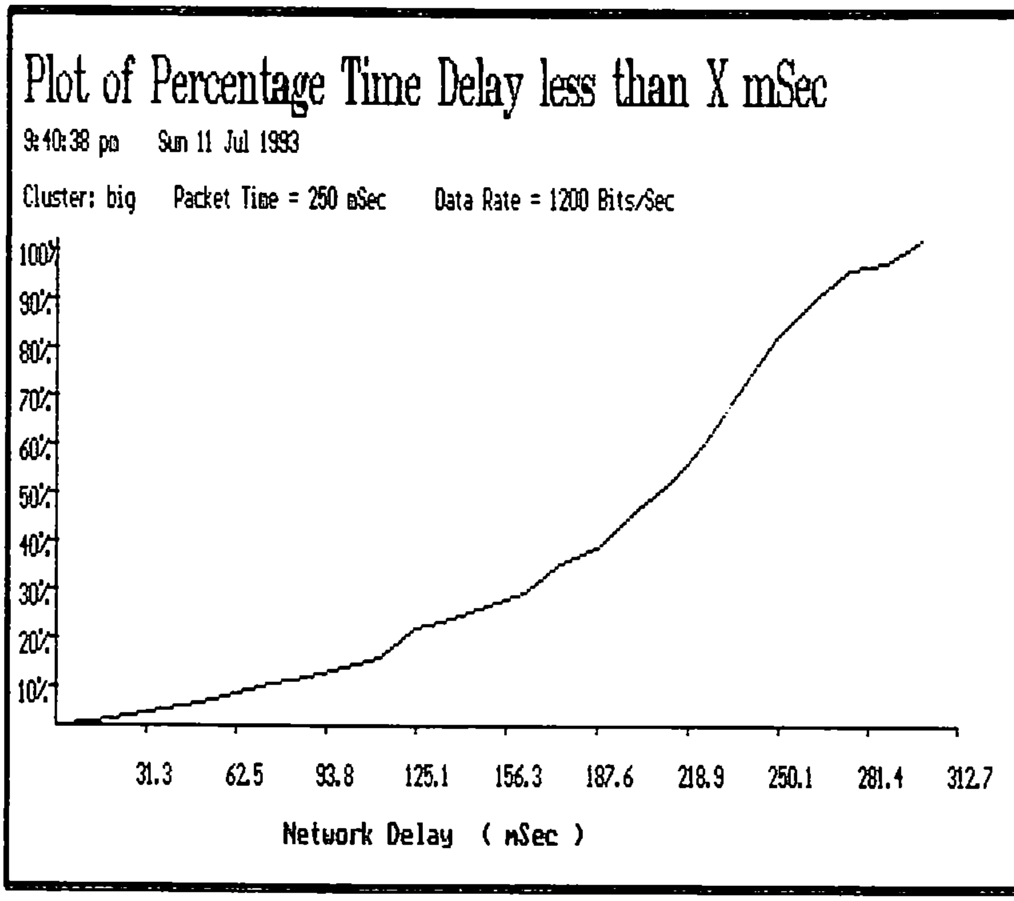
4. Ozekici S.

'Queuing Theory and Applications', Hemisphere Pub Corp, 1990.

5. Pierre S. & Hoang H.

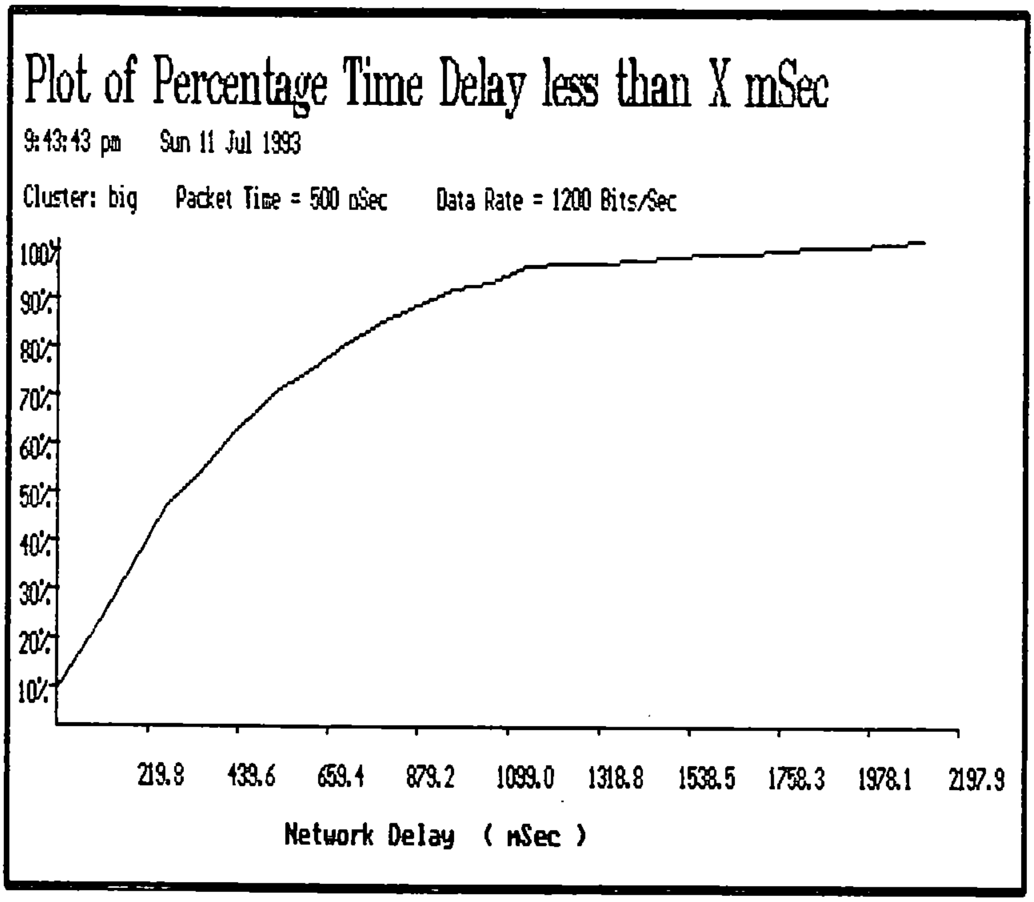
'An Artificial Intelligence Approach to Improving Computer Communications Network Topologies', J Ops Res Soc vol 41, no 5, 1990.

Result 1



Delay < 12.5 mSec for 0.4% of time Delay < 25.0 mSec for 1.3% of time
Delay < 37.5 mSec for 2.4% of time Delay < 50.0 mSec for 3.9% of time
Delay < 62.5 mSec for 5.1% of time Delay < 75.0 mSec for 6.8% of time
Delay < 87.5 mSec for 9.0% of time Delay < 100.1 mSec for 10.5% of time
Delay < 112.6 mSec for 12.4% of time Delay < 125.1 mSec for 14.2% of time
Delay < 137.6 mSec for 20.5% of time Delay < 150.1 mSec for 22.3% of time
Delay < 162.6 mSec for 25.1% of time Delay < 175.1 mSec for 27.5% of time
Delay < 187.6 mSec for 33.3% of time Delay < 200.1 mSec for 36.7% of time
Delay < 212.6 mSec for 44.6% of time Delay < 225.1 mSec for 50.5% of time
Delay < 237.6 mSec for 58.9% of time Delay < 250.1 mSec for 70.3% of time
Delay < 262.6 mSec for 80.7% of time Delay < 275.1 mSec for 88.2% of time
Delay < 287.6 mSec for 94.2% of time Delay < 300.2 mSec for 95.4% of time
Delay < 312.7 mSec for 100.0% of time

Result 2



Delay < 87.9 mSec for 7.4% of time Delay < 175.8 mSec for 19.8% of time
Delay < 263.8 mSec for 32.1% of time Delay < 351.7 mSec for 45.5% of time
Delay < 439.6 mSec for 52.8% of time Delay < 527.5 mSec for 61.7% of time
Delay < 615.4 mSec for 68.9% of time Delay < 703.3 mSec for 74.0% of time
Delay < 791.3 mSec for 78.7% of time Delay < 879.2 mSec for 84.0% of time
Delay < 967.1 mSec for 87.2% of time Delay < 1055.0 mSec for 90.3% of time
Delay < 1142.9 mSec for 92.0% of time Delay < 1230.8 mSec for 94.8% of time
Delay < 1318.8 mSec for 95.4% of time Delay < 1406.7 mSec for 96.0% of time
Delay < 1494.6 mSec for 96.5% of time Delay < 1582.5 mSec for 97.0% of time
Delay < 1670.4 mSec for 97.5% of time Delay < 1758.3 mSec for 98.0% of time
Delay < 1846.3 mSec for 98.4% of time Delay < 1934.2 mSec for 98.8% of time
Delay < 2022.1 mSec for 99.2% of time Delay < 2110.0 mSec for 99.6% of time
Delay < 2197.9 mSec for 100.0% of time