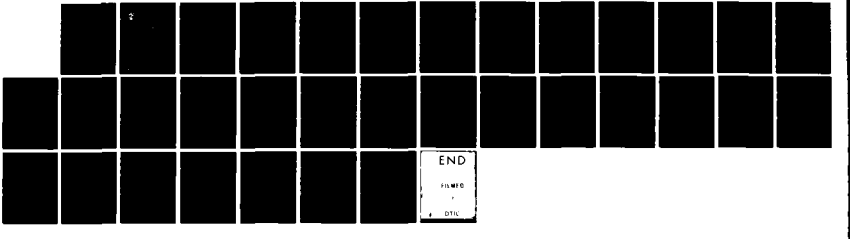
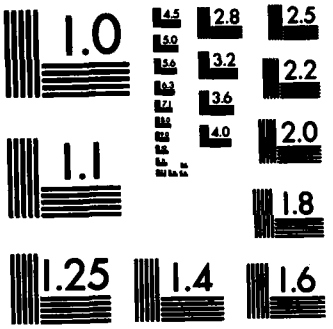
AD-A122 673 INTERNETWORKING IN THE MILITARY ENVIRONMENT(U) ROYAL SIGNALS AND RADAR ESTABLISHMENT MALVERN (ENGLAND) B H DAVIES ET AL. JUL 81 RSRE-MEMO-3391 DRIC-BR-80727 1/1

UNCLASSIFIED F/G 17/2 NL





UNLIMITED

RSRE
MEMORANDUM No. 3391

ROYAL SIGNALS & RADAR ESTABLISHMENT

INTERNETWORKING IN THE MILITARY ENVIRONMENT

Authors: B H Davies and A S Bates

PROCUREMENT EXECUTIVE,
MINISTRY OF DEFENCE,
RSRE MALVERN,
WORCS.

"THIS DOCUMENT IS THE PROPERTY OF HER BRITANNIC MAJESTY's GOVERNMENT and is issued for the information of such persons only as need to know its contents in the course of their official duties. Any person finding this document should hand it to a British Forces Unit or to a Police Station for its safe return to the MINISTRY OF DEFENCE (HQ security), LONDON SW1A 2HB, with particulars of how and where found. THE UNAUTHORISED RETENTION OR DESTRUCTION OF THE DOCUMENT IS AN OFFENCE UNDER THE OFFICIAL SECRETS ACTS OF 1911-39. (When released to persons outside Government Service, this document is issued on a personal basis and the recipient to whom it is entrusted in confidence, with the provision of the Official Secrets Acts 1911-39, is personally responsible for its safe custody and for seeing that its contents are disclosed only to authorised persons)."

UNLIMITED

ROYAL SIGNALS AND RADAR ESTABLISHMENT

Memorandum 3391

TITLE: INTERNETWORKING IN THE MILITARY ENVIRONMENT

AUTHORS: B H Davies & A S Bates

DATE: JULY 1981

SUMMARY

The increasing requirement for data communications in the military environment and the heterogeneous nature of the network technologies and protocols involved are highlighted. The main section of the paper discusses how the design of a military internet architecture is influenced by the military requirements especially that of survivability. Comparison with the civilian PTT approach to internetworking shows that while there are economic advantages to using civilian international standards where possible, these standards do not satisfy the military requirements. In particular the strategies for routing in a heavily damaged network environment and addressing hosts that migrate from one network to another must form an integral part of the overall architectural design. This results in gateways whose routing tables have a finer degree of detail of the internet topology than is usually required but which do not contain connection orientated information.

Finally, practical experience gained on the ARPA catenet system is described.

"Internetworking in the Military Environment"

I. INTRODUCTION

The increasing complexity and tempo of modern warfare has rapidly created the need for flexible data communications, parallel to those associated with the "information technology" growth in the civilian environment. The aim of this paper is to highlight the differences in emphasis between data communications in the civilian and military environments, and to examine the consequences of these differences. In particular, the importance of an overall communications architecture, in order to provide survivable and interoperable communications involving both present and future systems, cannot be overstated.

Experience gained in connecting a prototype military network to the ARPA catenet system and measurements made using internetworking data transport protocols are described. Enhancements to the system to improve survivablility and performance are suggested.

II. THE REQUIREMENT

To a large extent, the increase in the demand for data communications stems from the increasing use of computers, microprocessors and digital circuitry in weapons, sensor, and command and control systems. These devices are used for similar reasons to those pertaining in the civilian environment, in that they can perform well specified tasks faster, more reliably and more cheaply than human personnel. However, in order to accomplish the overall goal of efficient deployment of military resources, these geographically separated devices must communicate with each other and exchange information in a hostile environment. A distinctive property of the communications between these devices, is the very "bursty" or non-continuous nature of the information transfers, which makes packet switching an attractive means of providing the communications. In packet switching, bandwith is only allocated on demand, and therefore this techique allows considerably more efficient sharing of communication resources than the use of dedicated communication links. A further advantage of a well designed network, is the inherent survivablility of communciations that it provides. This is illustrated in figure 1a, which shows six computers interconnected by dedicated links. If any computer is to communicate with any other computer, then $n(n-1)/2$ links are needed, where n is the number of computers. If any one link fails, then two of the machines will fail to communicate. In figure 1c, which shows only nine links being used in a network, any two links may fail and possibly four, without completely cutting off communciations between the users of the network. This does not mean that networks in a damaged condition provide the same quaility of service as in their pristine condition, hence the necessity for priority markings to indicate which data is the most important. However, we can say that packet switching is an economical means of distributing the communciations resources in such a manner that it is difficult for the enemy to completely destroy communciations between users of the network.

So far we have described a single set of users connected to one network. However, there are many different types of networks based on different technologies and providing different types of service. This diversity of network types is due to the different user requirements and environments. For example, naval data communications may well be provided by a packet satellite network because of the large geographical

area of coverage required and the great mobility of the hosts or users of the network. In the forward area tactical environment, the data communcations may well be provided by a frequency hopping packet radio network, because of the extreme hostility of the electromagnetic environment. Finally, in an underground control centre, or on board a single ship, the communications may be provided by a "local area network". Figure 2 shows some of the key features of these different networks, from which it can be deduced that different control algorithms, error control strategies, routing etc are needed for the different networks. In particular, control and routing algorithms on single hop networks and broadcast networks will be very different from those on multi-hop networks.

Besides these different hardware technologies, the grade of service provided to the user may differ. For example, a network which is primarily designed for transporting sensor information, may well be optimized for providing minimum delay in the delivery of the data, rather than providing reliability of delivery, because of the perishable nature of the data. Thus, users who are primarily interested in reliable delivery would have to initiate transport control features on an end-to-end basis, to provide for loss and misordering of the data by the network.

There is a requirement for users on the different networks to communicate with each other [1]. In particular, the long haul communications may be provided by a common bearer network, which may interconnect forward area networks with local command centre networks. Also, with additional tasks and new capabilities, there will continue to be new and unknown data communications requirements, which will have to be integrated with existing systems.

The main requirements of data communications are that they should be secure, survivable and interoperable [2]. This paper concentrates on the survivability and interoperability issues, and the reader is referred to the references which concern computer and network security [3,4]. However, it is necessary to point out that the more interoperable the systems are, the greater the security risks, because there are more avenues of attack on the confidentiality and integrity of the data, by a greater number of personnel. In particular, "access controllers" or security sentinels in critical gateways, which interconnect networks, may restrict access to certain types of traffic, thus sacrificing survivablility and flexibility in the interests of security. Survivablility of communciations has many different meanings, but in its strictest sense it implies fully automatic routing around damaged switching components or links, and the ability to use alternate routes, even through other networks, in such a way that data integrity is maintained on an end-to-end basis.

III. REASONS FOR AN OVERALL ARCHITECTURE

To date, most communciations systems have not been designed with an overall communcations architecture in mind. This has resulted in great difficulty in providing interoperability with other systems. Because the modulation and coding, addressing and message representation, have often been combined, interconnection with another system has involved a very expensive interface box between the two systems. The disadvantages of this approach are:-

1) Each interface box is a special 'one off' design, which is custom built and therefore very expensive in design time and procurement cost.

2) Inevitably, in translating between one system and another, there will be certain features and sevices that will not have an equivalent in both systems.

3) Because of the processing power required to translate at all protocol levels, the interface unit will be a large and expensive piece of hardware. This has an effect on survivability, in that because the interface units are expensive, the minimum will be procured and the survivability of the overall communications will be determined by these vulnerable interface units.

The problem of deciding on the best architecture for computer to computer communications, has been the subject of sustained discussion over the passed decade. In particular, the International Standards Organization's subcommittee 16 has produced a major document in this field, "Reference Model of Open Systems Interconnection" [5]. The central thesis of this document is that the most flexible architecture is a layered one, in which each layer has a well specified function and provides a well specified service to the layer above it. In particular, any given layer views the layers below it as a single entity. This is analagous to structured programming, where the user of a procedure call is only interested in how parameters are passed to and from the procedure and not in the internal structure of the procedure. The seven layer model is illustrated in figure 3. Two points about the model are relevant to the discussion below. Firstly, the functional specification of each layer is more difficult to agree on, the higher the layer, because in these layers in the architecture there are more choices. Secondly, there has as yet been no ISO agreed protocols for implementing any of the layers. The model itself does not preclude more than one protocol implementing a given layer of the architecture.

IV. CURRENT STATE OF CIVILIAN STANDARDS

In Europe, with its highly regulated public communications authorities, there has been a very active co-operation among various countries to establish data communications standards from the outset. The CCITT (The International Telegraph and Telephone Consulative Committee), which is the corporate body representing the telecommunications authorities of these countries, has developed standard protocols, X25 [6], for levels 1,2 & 3 of the ISO reference model. It is important to note that in arriving at these standards, the PTTs (Public Telegraph and Telephone authorities) have identified that most customers want a connection orientated type of service, ensuring ordered and reliable delivery of packets. The network reserves the right, in event of a network error or congestion, to send a reset to both ends, indicating loss of data integrity. At present, no figures are available to indicate the frequency of such events. Because the main public networks in Europe are X25 networks, there has been considerable pressure on computer manufacturers to provide X25 hardware and software products off the shelf. This has lead manufacturers of private networks, in particular local area networks, to consider providing X25 accesses, in order to facilitate connections to existing machines and operating systems. Thus, X25 is rapidly becoming a de facto international standard in Europe.

What about the interconnection of X25 networks? Obviously, connecting networks which use the same access protocols and provide the same grade of service, is not so difficult a problem as interconnecting very dissimilar networks. Thus, there are X series protocols, X75, X121 [6], which enable PTT's to provide connections between users on different X25 networks, and although not all X25 facilites are available on internetwork connections, the service offered is analagous to STD dialling of international telephone calls. However, these protocols do rely on the X25 networks themselves, to route the internet packets to the gateways. It appears that private networks will not be allowed to connect to public networks via X75 gateways, and so gateways between private and public networks will have to provide a service between two X25 calls back-to-back, and will thus act as a staging post for the user's data.

Protocols for the transport layer (layer 4 of the OSI Reference model), are not so well developed as for the lower layers. However, in the United Kingdom a transport protocol [7] has been defined, and implementations above X25 have been realized. The most notable feature of this protocol is the flexible addressing structure, which allows connections to be established across different naming/addressing domains.

Before considering the applicability of these devlopments in the military environment, it is useful to consider some of the differences in emphasis, between civilian and military networks, and their usage.

V. COMPARISON BETWEEN CIVILIAN AND MILITARY NETWORKS AND THEIR USAGE

1) The usage of military networks in time of war is very difficult to predict. Although major exercises give some idea of the user demand, past experience has shown that these are slightly artificial and may not give a true picture. In civilian networks, usage can generally be accurately predicted by extrapolating present useage patterns, with economic and equipment sales factors being taken into account.

2) The availability of the full capacity of a military network may well be degraded when it is most needed, because links may be jammed and nodes and gateways physically destroyed. In the civilian environment, there is usually a very high availability of hardware and data links, with the use of standby power supplies and 'hot' spares for critical nodes such as gateways.

3) In general, there is a considerably higher degree of mobility of both users and networks in the military environment. In particular, airborne networks such as JTIDS (Joint Tactical Information Distribution System), with users such as fighter aircraft, will place stringent requirements on internetwork connections and survivability. A consequence of this will be that the users may well be completely unaware of the internet topology. While mobile access to networks will obviously develop in the civilian environment, in general it constitutes a fairly static community of networks and users.

4) One of the major advantages of geographically distributed databases, which are flexibly interconnected with communications links, is the decrease in vulnerability of the overall system to the total failure of a site (eg by physical destruction). Thus, when designing military networks it is important not to introduce an Achilles heel by, for example, employing a centralized network control centre. However, centralized control may well be the most convenient and cost-effective solution in civilian environment.

5) Both civilian and military network authorities wish to povide secure, survivable, interoperable, and guaranteed grades of service to their users. The questions arise as to how much the user is willing to pay for these properties, and how important the properties are. The question of the importance of the property, depends on the threats to the network, and these are obviously substantially greater in the military case. This means that the solutions for military networks may well be more expensive, in terms of implementation and running costs, than those for the civilian environment.

VI. TECHNIQUES FOR NETWORK INTERCONNECTION

At present there are two main architectural methods [8] for providing process to process communication across dissimilar networks. They are referred to as the "end-to-end" and "hop-by-hop" methods, because in the former, all the control information relevant to a particular data connection is held only in the source and destination hosts, while in the latter, connection orientated information is also held in various intermediate switching nodes, called gateways

The end-to-end approach is based on the assumption that all networks will offer at least an unreliable datagram service, i.e. if a sequence of packets is injected into the network then the destination will receive some of them, possibly misordered, and with possible duplication. Any improvement on this grade of service will be achieved by implementing end-to-end procedures to perform reordering, retransmission of losses and detection of duplicates. A legitimate criticism of this approach is that these upgrading procedures are acting across all the networks in the chain, which in the case of good networks means that there are extra overheads which involve needless expenditure. Thus, in the hop-by-hop approach, the required level of internet service is provided by procedures implemented across each network. This is obviously more expensive initially, in that the procedures are different for the different networks, but its running cost are cheaper because unnecessary control and retransmissions do not occur across the networks providing the higher grade of service.

There are also two schools of thought on addressing strategy, which are difficult to completely separate out from the ideas set out above. The first school, which has to date been associated with the end-to-end approach, is that all networks worldwide should have a unique network number allocated to it by a global authority. Thus, any host address can be uniquely defined worldwide by concatenating its network number with its host number. The addressing of internet packets is then simple. The other school believes that such international agreement on address formats is not achievable in the near future, and that there will exist multiple naming/addressing authorities. Thus, the address field will have to consist of a list of addresses in different formats, which will be parsed by the gateways of the differnt naming authorities as the packet wends its way through the internet system. This second system is considerably more flexible than the first, but as we shall see has other consequences as well.

To date, operational systems of the end-to-end variety have used a flat addressing space and the hop-by-hop systems have used the multiple domain system. A schematic representation of the protocol layering involved, in an internetwork connection across three networks, is shown for both the hop-by-hop and end-to-end approaches in figure 4. The hop-by-hop diagram clearly illustrates that the total service is provided by three concatenated services, involving different transport protocols on different types of networks. The end-to-end representation illustrates the singular nature of the transport service, which is independent of attributes of the underlying networks. We will now compare the advantages and disadvantages of the two systems, in the light of operation in the military environment.

1) Running Costs

The hop-by-hop approach has the advantage over the end-to-end approach as far as the civilian user is concerned, in that it is very 'tarrif' conscious (i.e. it only uses the minimum amount of transport protocol necessary to provide the required grade of service). Now as many of the European networks provide the high reliability of a virtual call service, this means that hop-by-hop implementations of the transport

service for these networks will involve minimum overheads in terms of extra bits to be transmitted, and therefore their running costs will be minimal.

In the end-to-end approach, every packet carries a full internet source and destination address in its header, so that it can make its own way to its destination. In the hop-by-hop approach, once the call has been set up, only the destination address for that particular network has to be carried, because the gateways on route contain addressing information for further hops.

2) Development Costs

The philosophy of the hop-by-hop approach implies a different protocol for each different type of network. This is not so serious in the civilian environment, because of the considerable influence of the CCITT standards, which means that most European public and private networks are of the X25 variety. Even local networks with very high speed interfaces are planning to implement an X25 access. However, in the military environment, where there is a considerably greater range of networks, this could require the development of a number of transport protocols.

3) Trusting Transit Networks

When a user makes a multi-net connection, using the hop-by-hop approach, it implies that he trusts the level of transport service being offered by the intermediate gateways in the internet route. Furthermore, it implies that he is happy with the reliability of intermediate gateways which ,albeit temporarily, take responsibility for his data at the termination of each hop. We believe that this is a state of affairs that is considerably more acceptable in the benign civilian environment than in the hostile military one.

In the end-to-end approach, only an unreliable datagram delivery service is expected from the set of concatenated networks, and loss of data in any intermediate switching node or gateway will be recovered by a retransmission from the source. Therefore, maintaining the bit integrity of the data transmission does not rely on the continuing correct operation of an intermediate node.

4) Addressing Strategy

In the multi-domain address strategy, if a user in one domain wishes to communicate with users in another domain, the user must know the topology of the interconnection of these domains, so that he can supply the information necessary for his data to reach the destination domain. This information could be obtained automatically for him, but it implies separate and possibly different bilateral aggreements between the various domain authorities.

In the end-to-end approach with a flat addressing space, each packet contains complete addressing information, and is free to find the best current route across all intermediate networks (figure 5). This dynamic internet routing has similar resource allocation advantages to dynamic routing on single networks. This flexibility of routing in the internet environment is more important in the context of the more rapidly changing scenario of the military environment.

5) Transport Control

The end-to-end control is certainly less flexible than the hop-by-hop control. Timeouts in particular, may vary by an order of magnitude, even on the networks in service today. End-to-end flow control, also requires more sophisticated strategies than are needed in the hop-by-hop method.

6) Gateway Complexity

One of the chief attractions of the end-to-end approach with flat addressing is the conceptual simplicity and relative smallness of the gateways with respect to the hop-by-hop approach. This is because the only modules that vary from gateway to gateway are the network access modules that pertain to each network (and these are just the modules needed on all host attached to that network). The fact that no connection orientated information is held in the gateway, greatly simplifies the action that the gateway has to take on receiving a packet and the amount of buffer storage it needs.

The above property ties in well with the gateway policy for military networks, namely that networks should be multiply connected by gateways in order to provide survivable internetwork communications. Thus, the "simplicity" of the gateways will result in cheapness and the ability to provide more than one gateway between every pair of networks.

Thus, although the end-to-end approach involves higher overheads in terms of packet headers, we believe that it offers considerably increased survivability in a hostile environment. Furthermore, in a situation in which users and networks are mobile, it is necessary for all networks to come under a single naming/addressing authority (e.g. NATO) if these changes in topology are to be distributed rapidly and efficiently throughout the internet system.

Taking the ARPA catenet (an interconnection of networks) system as a baseline model for a military internet system, we will briefly describe its salient features and then go on to discuss enhancements that will further increase the survivablility properties of such a scheme.

VII. THE ARPA CATENET SYSTEM

An example of the end-to-end approach with a flat address space, which has been running operationally for about 5 years, is the ARPA catenet system. This system connects about thirty different networks including land-line, satellite and radio based networks, as well as a variety of local area networks. The thinking and concepts involved in the architecture of this system have been fully described in a number of papers [9,10].

The protocols responsible for data transport in this system and their hierarchical relationship are shown in figure 6.

1) Internet Protocol (IP) [11]

This provides for transmitting blocks of data, called datagrams, from sources to destinations. Its main paramters are source and destination addresses which are globally unique. There are two main parts to the internet address, a network field (eight bits) and a host field (twenty four bits). Implementations of this protocol exist in the gateways and internet hosts. The datagrams are routed from one internet module to another through individual networks. In this approach, datagrams may be routed across networks whose maximum packet size is smaller than they are. In this case, a fragmentation module breaks up the packet into smaller packets, replicating enough information in the headers to allow reassembly at the destination. Reassembly does not take place in the gateways, because packets may take different routes to their destinations. There are a number of options available in the internet protocol and these are spcified in the control information of the header. Thus, the internet

header is of variable length.

2) Transmission Control Protocol (TCP) [12]

TCP is a data transport protocol appropriate to level 4 of the ISO reference model, and is especially designed for use on interconnected systems of networks. TCP is a connection oriented, end-to-end reliable protocol, designed to fit into a layered hierarchy of protocols which support multi-network applications. It provides for reliable interprocess communications, between pairs of processes in host computers, attached to distinct but interconnected computer communication networks. The TCP assumes it can obtain a simple, potentially unreliable, datagram service from the lower level protocols. It fits into a layered protocol architecture just above a basic Internet Protocol, which provides a way for the TCP to send and receive variable length segments of information enclosed in internet datagram envelopes. In order for the TCP to provide a reliable logical circuit between pairs of processes, on top of the less reliable internet communication system, it performs the functions of basic data transfer, data acknowledgement, flow control and multiplexing.

3) Gateway to Gateway Protocol (GGP) [13]

The gateway to gateway protocol is responsible both for distributing routing information through the gateways of the catenet and for advising communicating hosts of routing changes, congestion control and unreachable destinations. The basic routing algorithm, in use today, is the original ARPAnet routing algorithm. This involves gateways telling their nearest neighbours which networks they can reach and how many gateway to gateway hops are involved in the route. If a gateway is directly connected to a network, then it is said to be zero hops to that net. Gateways continuously monitor the state of the network access switch to which they are connected and their nearest neighbour gateways to ensure that routes through them are still available.

VIII. PRACTICAL EXPERIENCE OF THE ARPA CATENET SYSTEM

In the Autumn of 1978, RSRE set up a collaborative program of research and development in communications with the Advanced Research Projects Agency of the US Department of Defense. This collaborative program involved the connection of the PPSN (Pilot Packet Switched Network), our own in-house research network, to the ARPA catenet system, and providing terminal and file access from an internet host on the PPSN to some of the major Arpanet hosts. The first two years of the program were allocated to the development and implementation of a reliable connection between PPSN and the ARPA catenet system. During the early stages of the program, terminal access to the catenet system was provided by a Terminal Interface Unit (TIU), the major software modules, written in PDP-11 assembler code, were supplied by the Standford Research Institute [23]. Since then, we have implemented the DoD standard Transmission Control Protocol, the Internet Protocol, and the Gateway-to-Gateway Protocol in Coral-66. In addition, we have made several measurements on the performance of the catenet system, particulary in terms of round-trip delays, as the connectivity and the development of the catenet has evolved.

These measurements and experiences, along with those of other members of the internet working group, have highlighted some interesting problems, and as a consequence some enhancements to the Internet and Transmission Control Protocols have been proposed.

1) Implementation Issues

For a particular implementation of a protocol to operate as the designers of the protocol intended, it is necessary to not only specify the protocol, but to consider the implementation problems of that protocol. While each implementation will be 'tailored' to fit its own environment, there is a need for guidance on certain issues. In this section of the paper, a discussion of some of the questions which have arisen during our implementation of the internet protocols is given.

A) Internet Protocol Fragmentation and Reassembly

In the ARPA catenet scheme, all hosts must be prepared to accept datagrams of upto 576 octets in length, either in one piece or in fragments to be reassembled. To fragment a long internet datagram, the internet protocol module must create two or more new internet datagrams, and copy the internet header from the long datgram into each of the new datagrams. It must then divide up the data portion of the long datagram into parts, on a 8 octet boundary. The question arises as to how should this split of data be proportioned.

The first approach would be to divide the data portion into muliples of the maximum transmission size for the network to be traversed. This would produce the minimum number of resultant datagram fragments. However, the last fragment could be much smaller in size, depending upon the size of the original datagram.

An alternative approach would be to divide the original datagram up into near equal size fragments. This would produce the same number of fragments, as with the first approach, but the length of each fragment will be smaller in most cases. This has the advantage that these fragments will be less likely to be fragmented later by another gateway, and as a result is a better approach.

If the internet datagram contains options, then this involves the gateway in a significant amount of processing since not all of the options are copied on fragmentation. To simplify this procedure , we feel that a bit in the option type octet could be used to indicate whether the option should be copied on fragmentation.

On receiving a fragmented datagram, the IP module has to allocate the necessary resources for reassembly. This will include a buffer whose size is equal to the maximum reassembled packet length (576 octets), and control data structures to remember which fragments have been received, since only when the end fragment is received will the module know the total length of the reassembled datagram. If the module is unable to allocate these resources it must drop the datagram. The question that arises is should the IP module remember that it could not reassemble an incomming datagram, and when some reassembly resources become free, allocate them automatically for any future datagram with those attributes. This would be a more fairer way to allocate resources, but does involve the maintenance of connection information.

In order to provide a way of uniquely identifing fragments of a particular datagram, the IP header includes an identifier. It is appropiate in many cases, for the higher level protocols (e.g TCP) to choose the identifier, since the probability of delivery is improved if a retransmitted packet carries the same identifier as the original packet, since fragments of either datagrams can be used to construct the complete TCP segment. However, this requires the sending protocol to keep a table of identifiers, and in protocols like TCP this table would have to include the identifiers associated with all of the segments transmitted and which have not been acknowledged, for all TCP connections.

Internet datagram fragments are reassembled at the internet destination host. As datagram fragments can arrive in any order it is necessary for the IP module in the host to keep a record of the data octet blocks so far received. This usually takes the form of some type of bit mapping. Thus, whenever a fragment is received the corresponding bit in the map is set and a check is then done to see if the original datagram is fully assembled. Thus, packet reassembly involves a significant amount of processing.

Internet fragmentation does allow datagrams originating from networks with large packet sizes, to traverse other networks with smaller packet sizes, without enforcing any changes to the individual networks involved. However, the resultant problems should be noticed, and alternative solutions such as intranet fragmentation, or even reducing the size of the original datagram packets should be considered wherever possible.

B) Retransmissions in TCP

TCP can be used for communications over a variety of different networks, therefore there can be a wide variation in the round trip delays. As a consequence, a fixed retransmission period is not suitable, since in some cases there will be significant delays when a TCP segment is lost, while in others there will be unnecessary retransmissions, wasting the resources of the networks.

Round trip packet delays between different hosts on the current catenet have been measured to be in general between 1/2 and 8 seconds. These measurements have also shown a significant variation in the delays during a connection to a remote host. As a consequence, we have implemented a dynamic timeout algorithm for retransmissions in TCP.

This algorithm measures the time elapsed between sending a data octet with a particular sequence number, and receiving an acknowledgement (ack) that covers that sequence number (thus one does not have to match sends and acks one for one). Using that measured elapsed time as the round trip time (RTT), we compute a smoothed round trip time (SRTT) as:

SRTT = (ALPHA * SRTT) + ((1 - ALPHA) * RTT)

and based on this, compute the retransmission timeout (RTO) as:

RTO = min{BOUND, BETA * SRTT}

where BOUND is an upper bound on the timeout (e.g., 30 seconds), ALPHA is a smoothing factor (e.g. 0.9), and BETA is a delay varience factor (e.g. 1.5).

The performance of this algorithm has been very good and has significantly reduced the number of retransmissions.

C) Silly Window Syndrome in TCP

TCP provides a means for the receiver to govern the amount of data sent by the sender. This is achieved by returning a "window" which indicates how many more octets of data, beyond that acknowledged by the packet, that the receiver is willing to accept. The silly window syndrome arises where a TCP has a lot of data to send (e.g. in file transfers) and results in many small segments being sent. The receiver reports an increase in window size each time a small amount of data has been processed, and the sender immediately sends a new segment to fit that additional window.

One way to prevent this, is for the receiver not to report a new window unless the increase is a reasonable size. The receiver can acknowledge incoming segments at any time, but limit window updating to points when a reasonable increase can be made. The sender can also try to prevent this by only sending big segments and waiting until the window is large enough to allow it. If long delays are involved, then it may be acceptable for the sender to exceed the window he has received, and hope that by the time the remote end receives the TCP segment then its window would have increased.

If the receive window is zero, then it is up to the sender to probe the remote end with a at least one data octet. Clark [18] suggests that it may be better to send an octet of old data. To minimise waste of the network resources, the remote end could send out an ACK packet when its window is increased significantly from zero. Because the acknowledgement containing the new window size could get lost, the sender should poll the remote end with packets containing several new octets of data.

2) The RSRE Connection to the ARPA Catenet

During the past 2 years, the RSRE connection to the ARPA catenet has changed significantly [22]. The current configuration is shown in figure 7. The RSRE internet host (PDP-11/23) contains the standard internet protocols of Telnet,TCP and IP all writen in the high-level language Coral-66, and which run under our own virtual memory operating system EMMOS [20]. The link level protocol, X25 level 2, is used to interface to the PPSN. This protocol is implemented on a microprocessor communication interface (X25 line unit) which is connected to PDP-11 hosts via a standard interface [19,21].

The PPSN is connected to the rest of the catenet via the RSRE gateway. The gateway is also a PDP-11/23 micro-computer, with all of the protocol code written in Coral-66. The gateway has three network interfaces on it, each using a X25 line unit. They are used to provide, 1) access to PPSN, 2) a test port which can be directly connected to a measurement host, and 3) an interface which connects the RSRE gateway to a gateway at University College, London (UCL) via a 9.6k bits/s Post Office line.

The UCL gateway is connected to two other networks, 1) UCL net and 2) Satnet (ARPA packet satellite network). The connection to Satnet is via the Goonhilly SIMP (Satellite Interface Message Processor). Packets destined for Arpanet are forwarded by the Goonhilly SIMP, over the shared 64k bits/s half duplex satellite channel to the Etam SIMP, and from there they are forwarded on to the BBN gateway, and hence into Arpanet.

2) Catenet Measurement Performance

For the purpose of interactive traffic, the delay in the catenet is important. At RSRE, we have made a significant number of measurements on round trip delays in the catenet, and these measurements have showed where some of the problems in the system were located.

The measurements were made using a traffic generator which could output either 1 data byte packets (as usually found in interactive traffic to remote echoing hosts) or 128 data byte packets (as may be used when files are listed or being transfered between host). The gateways in the catenet, and the SIMP's of Satnet, have datagram echo facilities built into them. In addition, a program has been written by UCL which when run on a TOPS-20 computer will also echo datagrams.

To measure delays in the catenet, the measurement host at RSRE stamps the local time into the out-going packet, just below the internet header. On receiving the echoed packet back, the time stamped in it is compared with the current local time and

the round trip delay determined. These delays will not only include the network transition times, but also any internal delays in the gateways and hosts.

The single round trip delays for the RSRE, UCL and BBN gateways are shown in figure 8. A total of 1000 internet datagram packets transmitted by the generator, at a rate of 1 packet/sec were 'fired' at each of the gateways. Each packet carried 1 ASCII character, in addition to the 6 bytes needed for time-stamping. The results for the RSRE and UCL gateways are what we would expect from theory, based on 9.6k baud lines. The results for the BBN gateway show a peak at 2 sec. This is in agreement with the expected Satnet performance. The secondary peak in the BBN histogram, is possibly due to a higher level of retransmissions across Satnet. In figure 9 the results for the Goonhilly and Etam SIMPs are shown. The additional delay, by echoing off the Goonhilly SIMP, rather than the UCL gateway, is very small due to the 48k baud line which interconnects them. The results for the Etam SIMP agree with those for the BBN gateway, with the exception of the secondary histogram peak, and indicates that the delay and dispersion is probably due to Satnet. Figures 10,11 shows the results for the SRI-PR1 and NDRE gateways respectively. The effect of echoing packets off the SRI-PR1 gateway, which is on the far side of Arpanet, rather than the BBN gateway is to add an additional time delay of approximately 0.5 sec. to the single round trip. The measurements for the NDRE gateway show a surprising longer delay than expected. This could possibly be due to a higher error rate being received at the Tanum simp, resulting in a significant number of retransmissions across Satnet. Using the internet echo program on ISIE, we were able to measure the single round trip delay to our mailbox host. These results are shown in figure 12, and are very similar to those for the SRI-PR1 gateway.

IX. ENHANCEMENTS TO THE CATENET SYSTEM

There are a number of situations, peculiar to the military context, which are not catered for by the algorithms presently used in the catenet. Before discussing these and possible enhancements to the catenet which would improve its survivablility in the military environment, we must introduce the concepts of "partitioned networks" and "source routing".

A "partitioned network" is one that is so badly damaged that there exist no paths between certain of its switching nodes. Typically, this results in two or more subsets or partitions of nodes, within which communications are possible, but which cannot communicate with each other. Hosts connected to different partitions cannot communicate in the usual way. However, if this network is connected by more than one gateway to the catenet system and there is at least one gateway on each partition, hosts could still communicate by an internetwork path as illustrated in figure 13. The concepts of routing to partitioned networks are concerned with automatic and efficient routing of packets under the conditions mentioned above.

The principle of "source routing" is one of providing some of the routing intelligence in the packet header, by providing not just the destination address, but also some or all of the intermediate node addresses through which the packet has to pass. This facility is provided as an option in the present DoD Internet Protocol.

1) Changes to the Catenet Routing Algorithm

The catenet system, as presently configured, permits routing around damaged networks and gateways. It assumes that hosts know the addresses of their local gateways, and are prepared to poll these gateways to determine their status, and have procedures for using alternate gateways, if the primary one is congested or

inoperative. Presently, routing to a partitioned network would involve knowing the topology of the catenet and inserting the routing information in the packet header in the form of a source route. This is perfectly feasible, but in a fast changing military environment it would be preferable if the gateways contained enough information to perform automatic routing to hosts on partitioned networks.

If the internet system of gateways is regarded as a super-datagram network, whose node to node protocol is the Internet Protocol, then it would seem reasonable that the internode routing be based on gateway or node identifiers. The routing information distributed to gateways should permit routing to a specific gateway, rather than to a network. As there may be more gateways than networks, this will involve the storage of more information in the gateways than at present. However, if there are additional gateway nodes for providing survivability, it is a waste of resources if the information is not disseminated and used when most needed.

There are two reasons for wishing to change the present catenet routing algorithm:-

(i) The present algorithm suffers from oscillations when certain link failures occur, because it uses repeated minimization to compute the shortest path. Presently, this problem is overcome by having a narrow range of link costs.

(ii) The granularity, or fineness, of the information distributed by the present algorithm which performs routing to networks, is insufficient for automatic routing to partitioned networks. This is because the route into a destination net via two different gateways may be wildly separated, as illustrated in figure 14. If the network is partitioned, we need to specify the entry into the net rather than just the net.

A recognized candidate for the improved routing algorithm is a modification of the New Arpanet Routing Algorithm [14], which is currently used on Arpanet. Using this algorithm, all the gateways broadcast information to all other gateways using a flooding technique. In particular, two types of information are disseminated:-

(i) Each gateway broadcasts the names of the nets to which it is directly connected.

(ii) Each gateway broadcasts the names of its neighbours with which it can communicate.

From this information, all gateways can determine which networks are partitioned, because a partitioned net will have two or more gateways attached to it which are unable to communicate. Having implemented this algorithm, there are one or two additional techniques that are necessary for dealing with routing to partitioned networks. The main remaining problems are, determining the partition in which the destination host is located, and specifying this in packets to be sent to that host. Now specifying the partition could be accomplished by specifying the identifier of the gateway through which the partition communicates with the rest of the catenet. However, at present there is no format for specifying gateway identifiers in the internet header. The determination of which partition the destination host is in, is best done by the gateway connected to the source host's network. This gateway will know how many partitions the destination network is divided into, and the entry gateways to these partitions. When the connection is being set up, the opening packet will be sent to all partitions, and the resultant reply will contain the relevent partition identifier. A minor expansion of the internet header will be required for specifying gateway identifiers in the internet packet headers.

2) Mobile Hosts in the Military Environment

There are already a number of requirements for aircraft flying from one tactical net to another, to be able to maintain communication with a ground based command and control centre [15]. There has been considerable discussion on posssible solutions to this problem [16,17]. The solution should, if possible, avoid using a centralized database. The disadvantage of using a centralized database in the military environment is not only its vulnerability, but also that a separate communication must be successfully performed with the database, as a prerequisite for a successful connection to the mobile host. Furthermore, as the host moves from one net to another, updates to the data base must be made in a timely manner. Obviously, a third party has to be involved if two mobile hosts wish to communicate. However, the ground control centre is a natural anchor for mobile communications, and if the TCP connection identifiers were divorced from physical addresses, the scheme below would provide total data integrity as the mobile host changed networks.

An interesting point, that is immediately highlighted when considering this problem, is that the unique identification of a TCP connection is at present tied down to physical addresses. We believe that this is undesirable, and has led to the present restricted attempts at solving this problem. We believe that unique TCP identifiers should be exchanged at the start of the connection and that these be used throughout, so that any changes in the physical addresses can be exchanged without closing the connection (i.e. when the aircraft changes nets it inserts its new address in the source address field, this is then used by the ground to continue the connection). It is possible that there will be a little hiccup as the change over from one net to another occurs, because packets may arrive out of order, however retransmissions would take care of this. It would obviously be the responsibility of the mobile host to 'login' to the ground centre on entering a network, so that a connection could be opened up from the ground. An alternative approach would be to include another protocol layer directly above the TCP layer. This new protocol would be responsible for opening and closing TCP connections and maintaining data integrity as the mobile host moved onto another network. The disadvantage of this approach, is the necessity to transfer the mobile host's new address on a three way handshake basis, before the host moved onto the new network.

3) Congestion Control in the Catenet

The catenet is essentially a super datagram network, and congestion control consists of using all possible routes to the best advantage and being able to offer a graceful degradation of service when the users demands exceed the network resources. It is important that fairness is exercised in providing a service to users, assuming that they are of the same priority. The above implies that the cost of a route should change if substantial queues build up on it, so that alternate routes become preferable in an SPF (shortest path first) routing algorithm. The change in cost will be reflected in the routing updates, and alternate less congested routes will be prefered. This requires a more realistic measure of internet routing costs, than the number of gateway to gateway hops used at present. This needs to be implemented on the catenet for realistic trials, even though the numbers of alternate routes is very small. Having thus made the best use of the internet resources, the only remaining action is to throttle off users when, by their weight of numbers, they overload the system. This throttling must be fair, bearing in mind priorities. One aspect of the fairness problem is that gateways handle packets on an independent datagram basis and are not therefore conscious of "greedy" users disobeying advisory flow control messages. A full solution of this problem would require a complex control theory model to be solved. This would involve the knowledge of the queuing sizes and delays on all intergateway links. The despatching of packets from the initial gateway would only occur when its journey through the system could be undertaken without it exceeding a

specified delay band.

X. SUMMARY

The main differences in emphasis between internetworking in the civilian and military environments have been described. In particular, the survivability requirements of the military users and the mobility of some of them, make the hop-by-hop technique employed by civilian networks totally unsuitable. The ARPA catenet approach provides the basis for a military internetwork system, but several enhancements are needed to meet all requirements.

X1. ACKNOWLEDGEMENTS

Many of the concepts presented in this paper have been widely discussed in the ARPA internet community. The authors wish to thank their colleagues in the internet community for their enthusiastic co-operation in the program of measurements described in section VIII.

XII. REFERENCES

[1]. G.E. LaVean, "Interoperability in Defense Communications", IEEE Trans Comm, vol com-28, no 9, pp1445-1455, Sept 80.

[2] F.F. Kuo, "Defense Packet Switching Networks in the US", Interlinking of Computer Networks,pp307-313, NATO ASI, Bonas Sept 78.

[3] R.B. Stillman & C.R. Defiore, "Computer Security and Networking Protocols", IEEE Trans Comm, vol com-28, no 9, pp1472-1477, Sept 80.

[4] D.H. Barnes, "Provision of End-to-end Security for User Data on an Experimental Packet Switch Network" IEE 4th Intnl Conf on Software Engineering for Telecommunications Switching Systems, Warwick July 81.

[5] Reference Model of Open Systems Interconnection, ISO/TC97/SC16/N227, International Standards Organization 1979.

[6] CCITT Recommendations X Series "Public Data Networks", Orange Book, ITU, Nov 80.

[7] British Post Office User Forum, "A Network Independent Transport Service", Feb 80.

[8] J.B. Postel, "Internet Protocol Approaches" , IEEE Trans Comm, vol comm-28, no 4, pp604-611, April 80.

[9] V. Cerf & R. Kahn, "A Protocol for Packet Network Interconnection",

Comput. Networks, vol 3, pp259-266, Sept 74.

[10] V. Cerf "DARPA Activities in Packet Network Interconnection", Interlinking of Computer Networks, pp287-313, NATO ASI, Bonas, Sept 78.

[11] DARPA, "DOD Standard Internet Protocol", IEN-128, Defense Advanced Research Projects Agency, Jan 80.

[12] DARPA, "DOD Standard Transmission Control Protocol", IEN-129, Defense Advanced Research Projects Agency, Jan 80.

[13] V. Strazisar, "How to Build a Gateway", Internet Experiment Note 109, Aug 79.

[14] J.M. McQuillan, I. Richer & E.C. Rosen, "The New Routing Algorithm for the ARPAnet" IEEE Trans Comm, vol comm-28, no 5, pp711-719, May 80.

[15] V.G. Cerf, "Internet Addressing and Naming in the Tactical Environment" Internet Experiment Note 110, Aug 79.

[16] C.A. Sunshine & J.B. Postel, "Adressing Mobile Hosts in the ARPA Internet Environment Internet Experimental Note 135, March 80.

[17] R. Perlman, "Flying Packet Radios and Network Partitions", Internet Experiment Note 146, June 80.

[18] J. Postel, "Internet Meeting Notes - January 1981" Internet Experiment Note 175, March 1981.

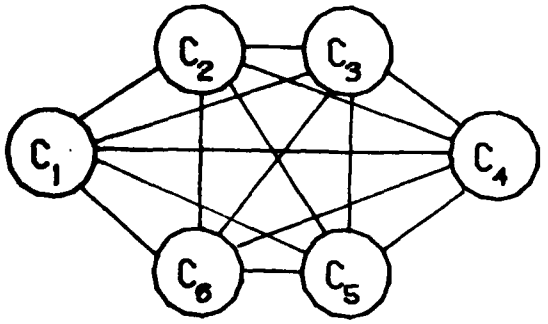
[19] A.F. Martin & J.K. Parks, "Intelligent X25 Level 2 Line Units for Switching", Data Networks: Development and Uses, Online Publications Ltd., pp371-384, 1980.

[20] S.R. Wiseman & b.H. Davies, "Memory Management Extensions to the SRI Micro Operating System for PDP-11/23/34/35/40", Internet Experiment Note 136, May 1980.

[21] A.F. Martin & J.K. Parks, "Experiences of implementing X25 Level 2 Packet-Switched Protocol", The Impact of New LSI Techniques on Communications Systems, IEE Colloquium Digest 1980/41, Oct. 1980.

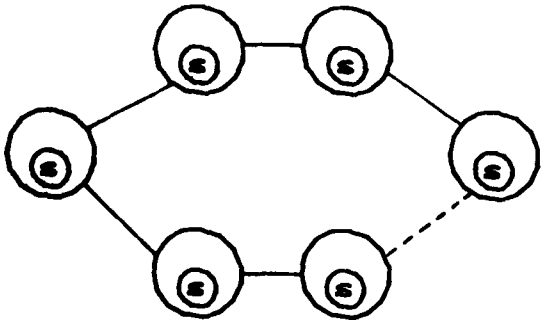
[22] B.H. Davies & A.S. Bates, "Internetworking in Packet Switched Communications: First Report on the RSRE-ARPA Collaborative Program", RSRE Memorandum No. 3281, July 1980.

[23] J. Mathis, "Micro Operating System Notebook", Stanford Research Institute, 1978.



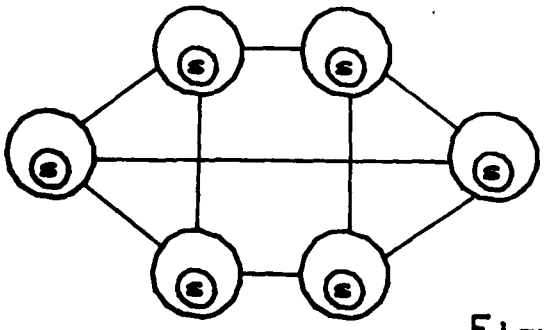
(a)

6 computers with 15 non-switched communication links. To allow for any one link failing there would have to be 30 links.



(b)

With integral switching software (S) (or front end processors) only 5 links are necessary



(c)

With 6 links any single link can be destroyed without cutting off communications.

With 9 links any two can be destroyed.

Figure 1

DIFFERENT PACKET-SWITCH TECHNOLOGIES

NETWORK TYPE / PROPERTY	FIXED TERRESTRIAL NETWORK	PACKET SATELLITE	PACKET RADIO	LOCAL AREA NETWORK
SWITCHING MODE	multihop	broadcast single hop	broadcast multihop	broadcast single hop
AREA OF COVERAGE	10-1000 Km	global 0-20,000 K	10 Km	0-10 Km
LINK BANDWIDTH	9.6-64 Kbits/sec	2.4 Kbit - 2 Mbits/sec	16-100 Kbits/sec	1-50 Mbits/sec
MOBILITY	fixed (possible mobile access)	mobile	highly mobile	fixed
ERROR RATE without coding	10^{-6}	10^{-5}	10^{-3} - 10^{-5}	10^{-9}
LINK COMMS	radio relay land line	radio/spot beam	omni-directional radio	coax twisted pair optical fibre
VULNERABILITY	low	high	medium	very low
COST	medium	high	high	very low

FIGURE 2

ISO MODEL FOR OPEN SYSTEM INTERCONNECTION

Application Layer	User Programs or Processes that wish to exchange information
Presentation Layer	This layer is concerned with data formats of information exchanged
Session Layer	This layer is concerned with synchronization and delimiting of information exchanges
Transport Layer	This layer provides a universal data transport layer independent of the underlying network
Network Layer	This layer provides network access and routing
Link Layer	This layer provides data transmission over a potentially unreliable link
Physical Layer	This layer specifies electrical signalling for data and control of the physical medium

FIGURE 3

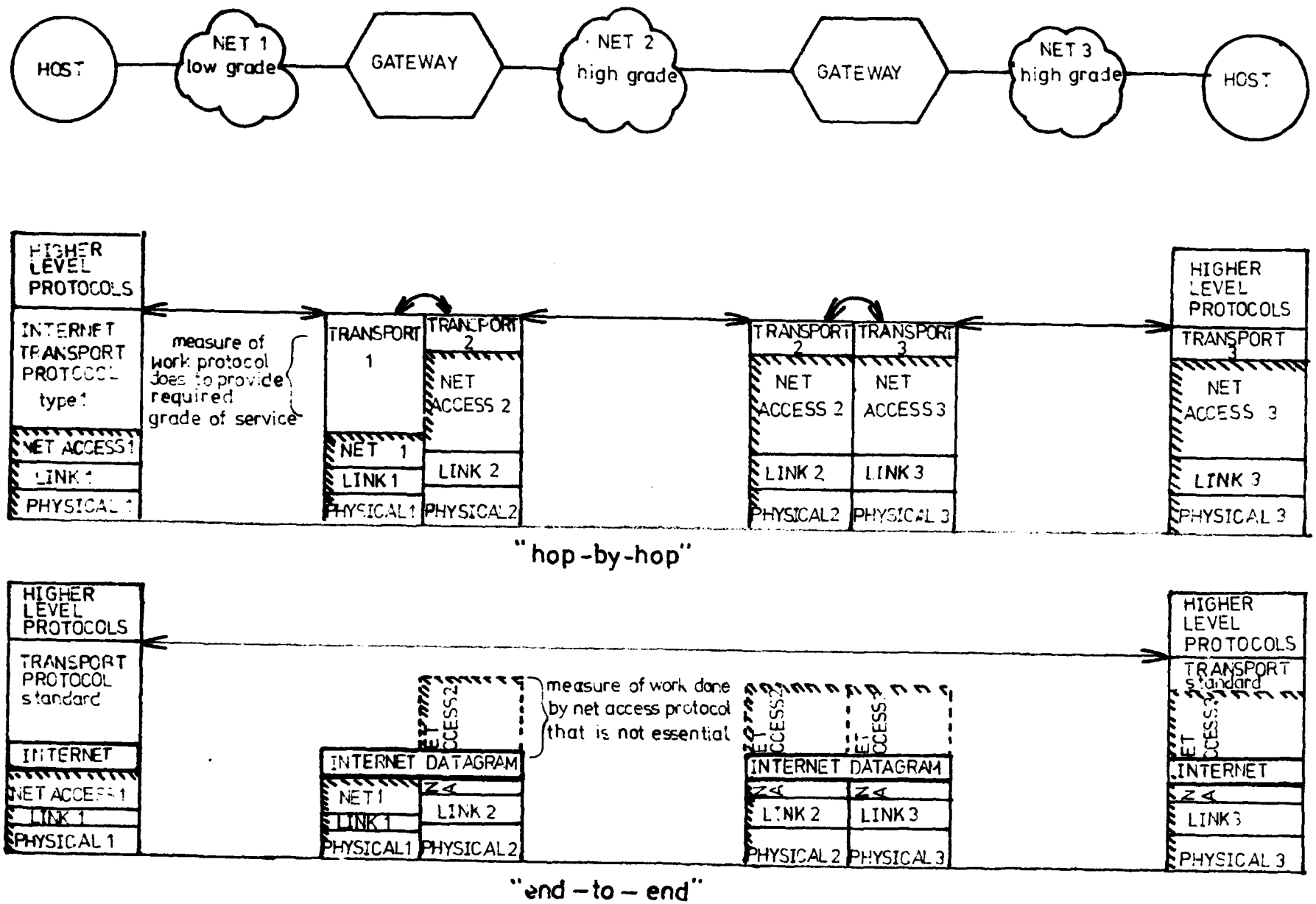
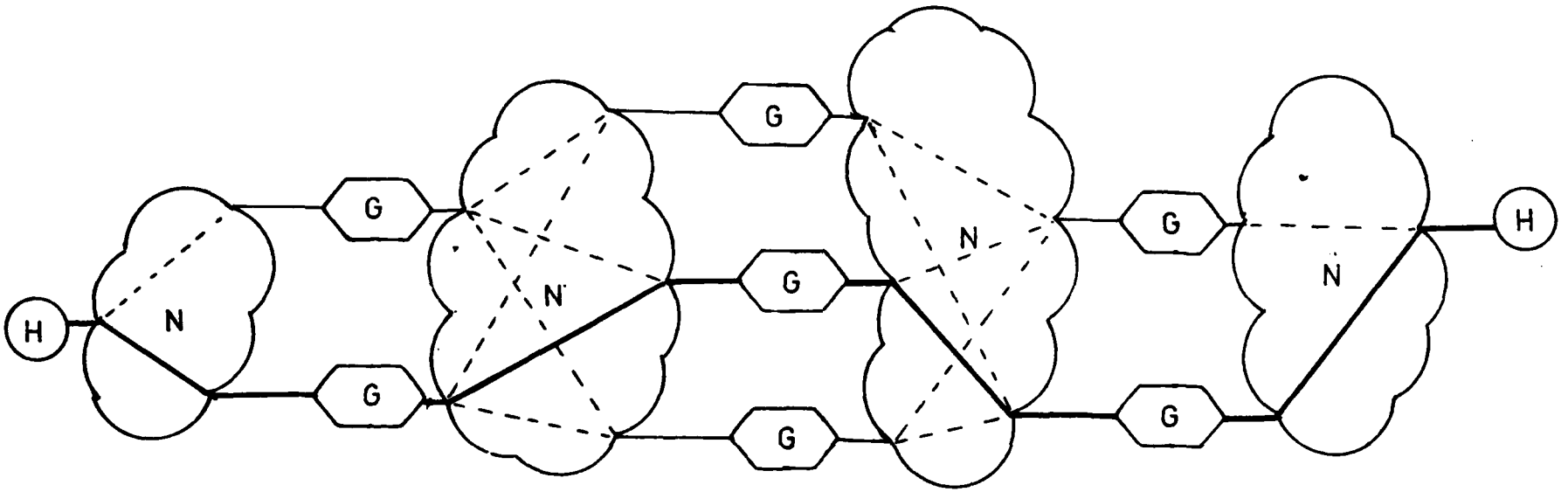


FIG.4 REPRESENTATION OF HOP-BY-HOP & END-TO-END PHILOSOPHIES OF INTERNETWORKING



internet datagrams in the "end-to-end" approach may take any of the dashed or the solid routes, but data in the "hop-by-hop" approach takes the solid route set up when the connection was established.

FIG.5 ROUTING FLEXIBILITY OF END-TO-END APPROACH

PROTOCOL LAYERS IN INTERNETWORKING

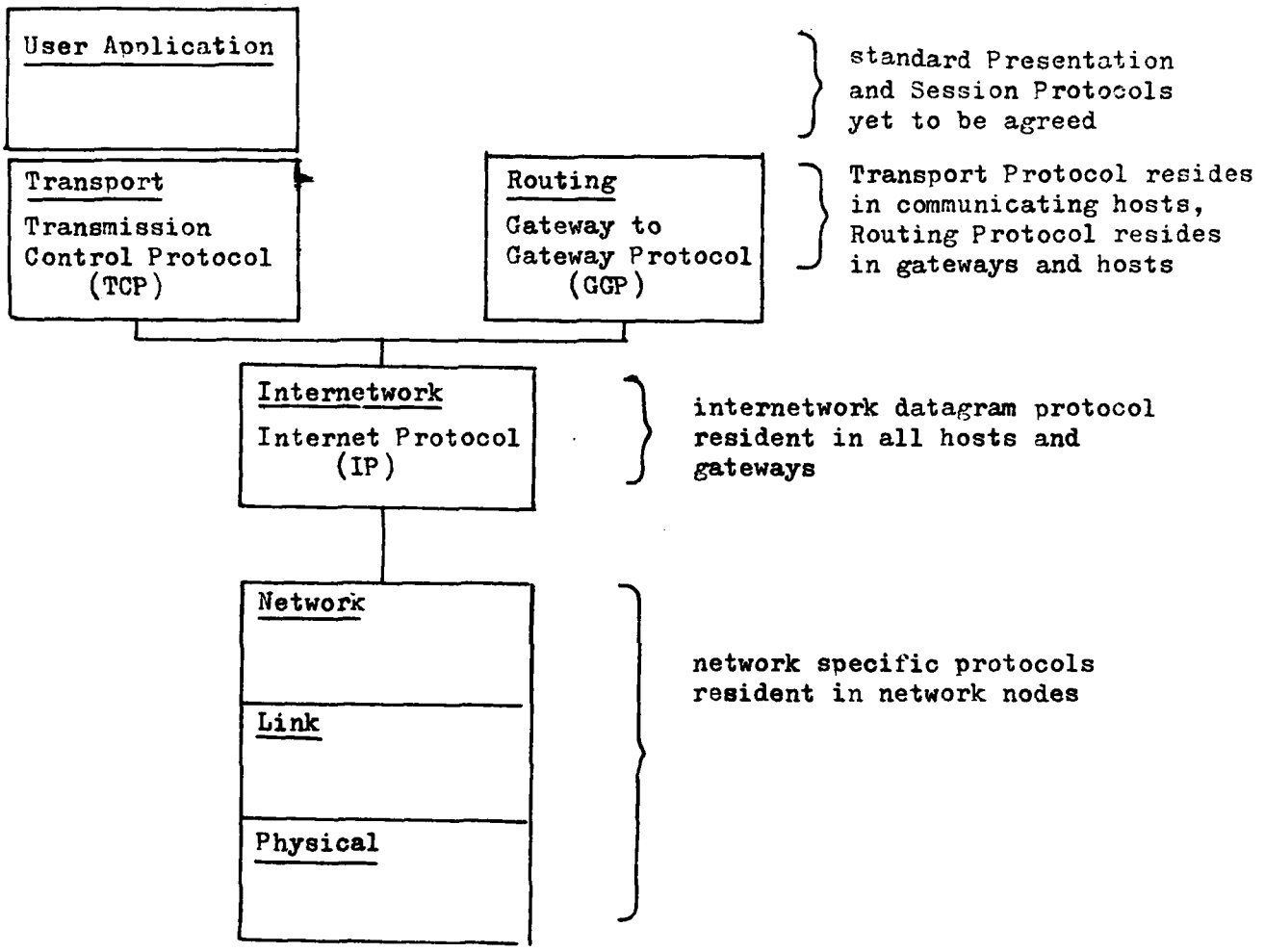


FIGURE 6

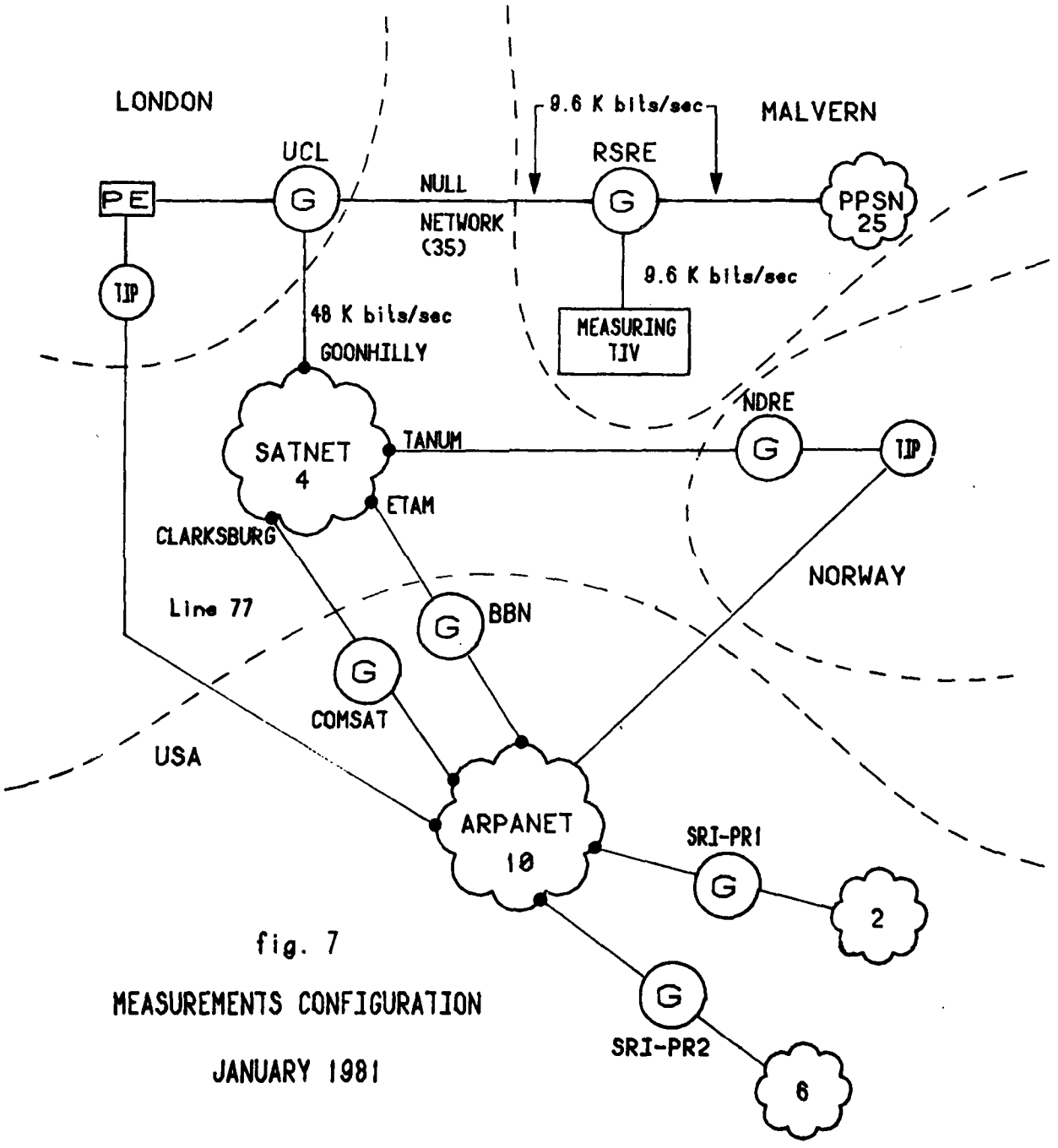


fig. 7

MEASUREMENTS CONFIGURATION

JANUARY 1981

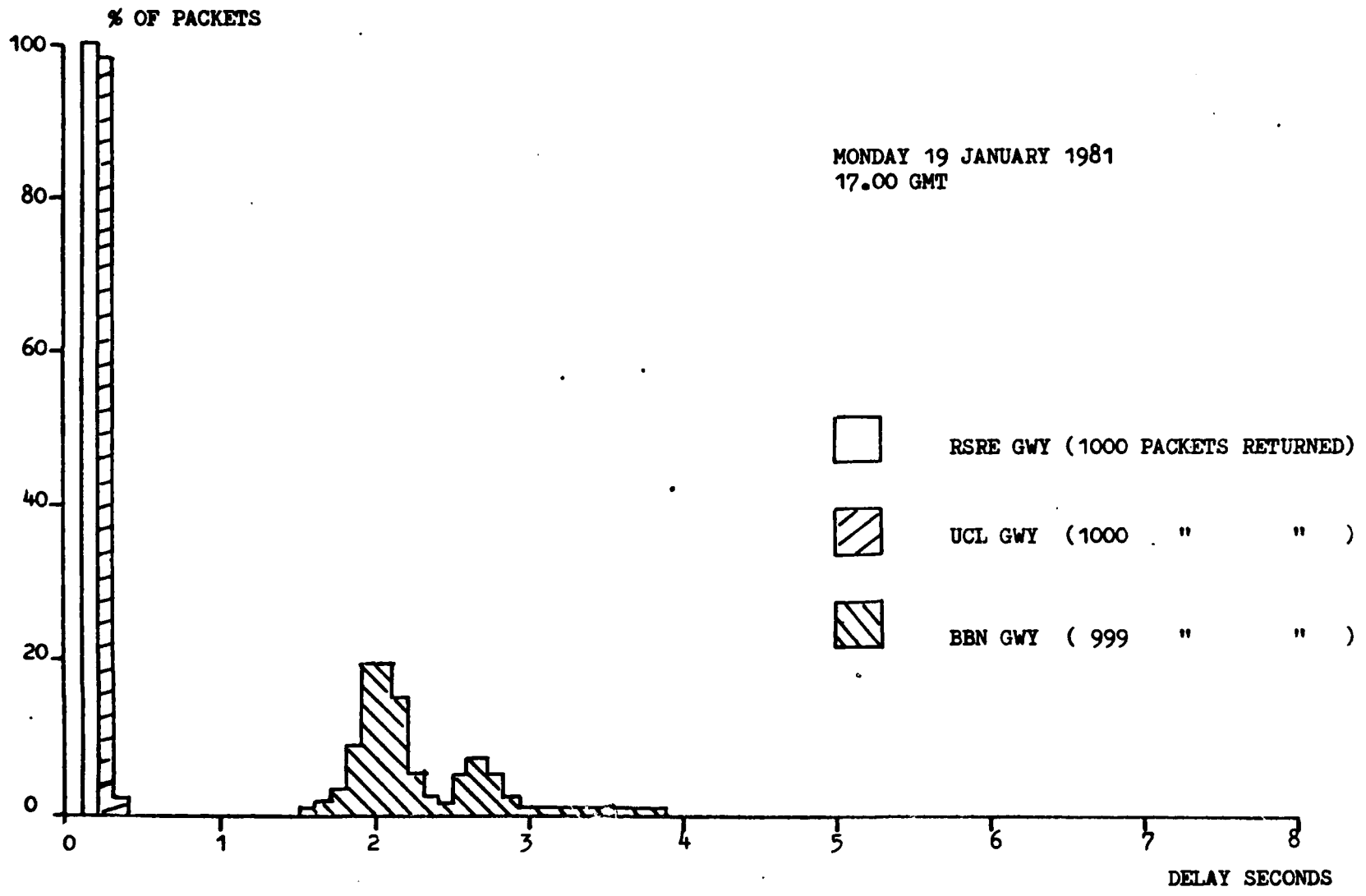


FIG. 8. SINGLE ROUND TRIP DELAY MEASUREMENT

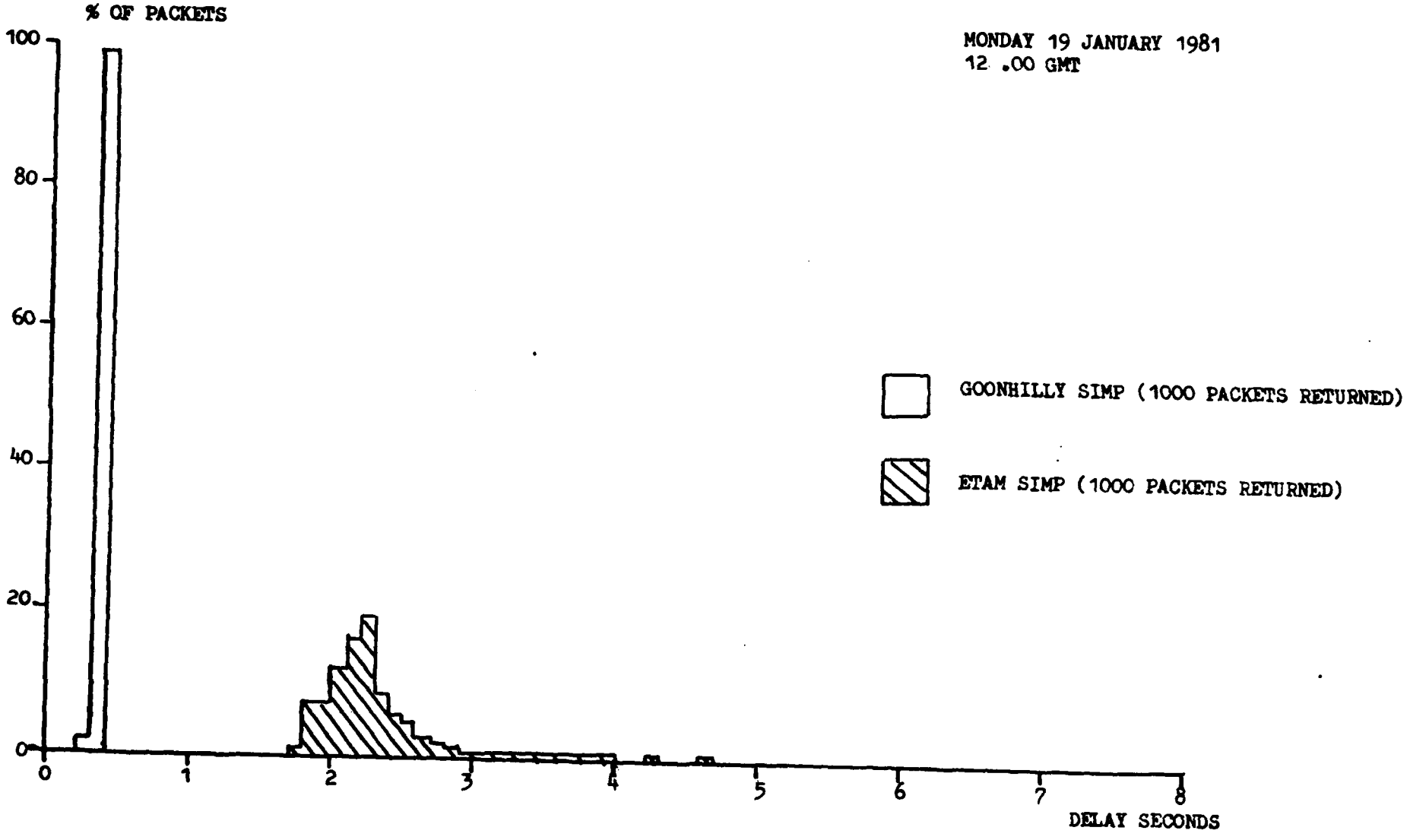


FIG 9. SINGLE ROUND TRIP DELAY MEASUREMENTS

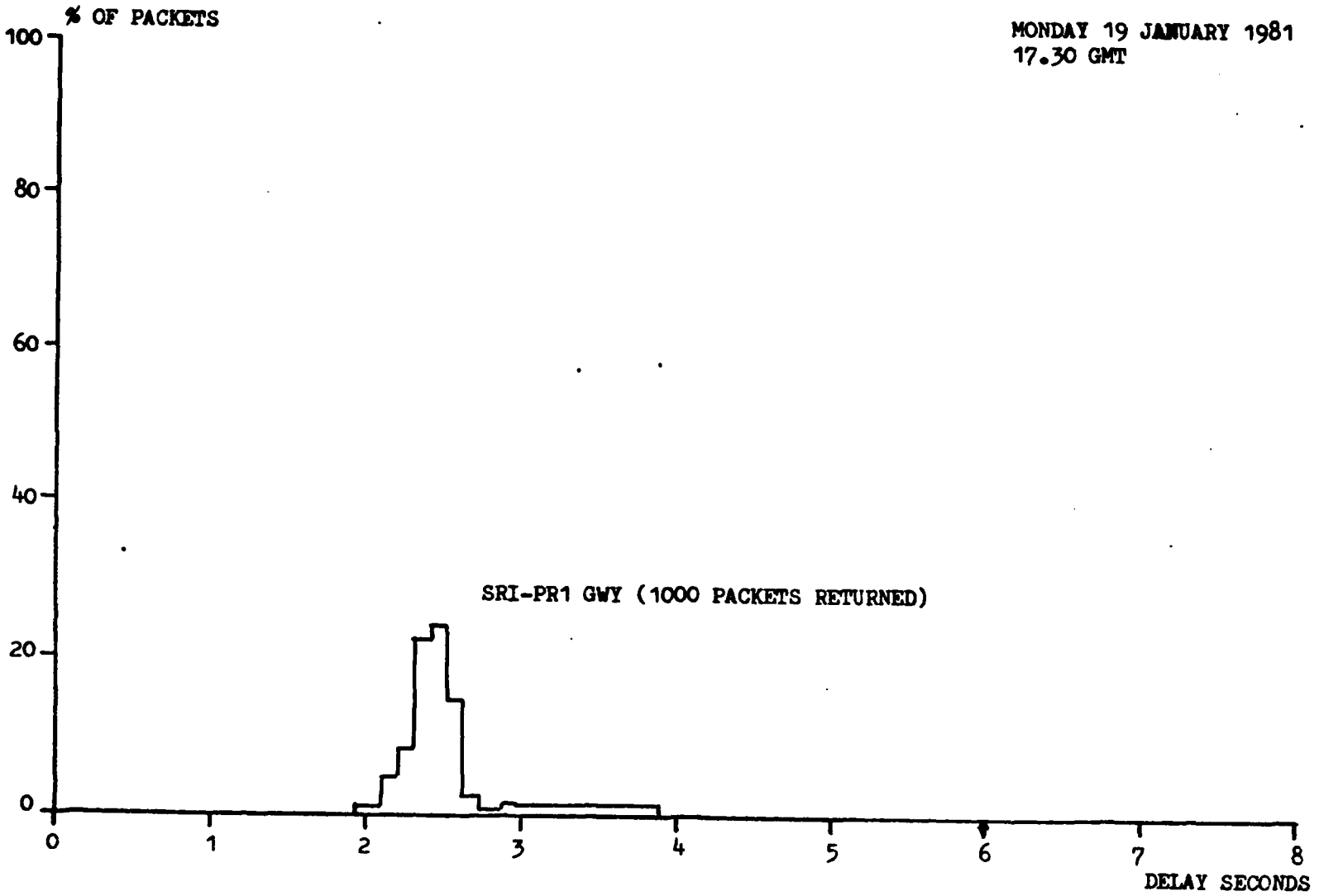


FIG 10. SINGLE ROUND TRIP DELAY MEASUREMENTS

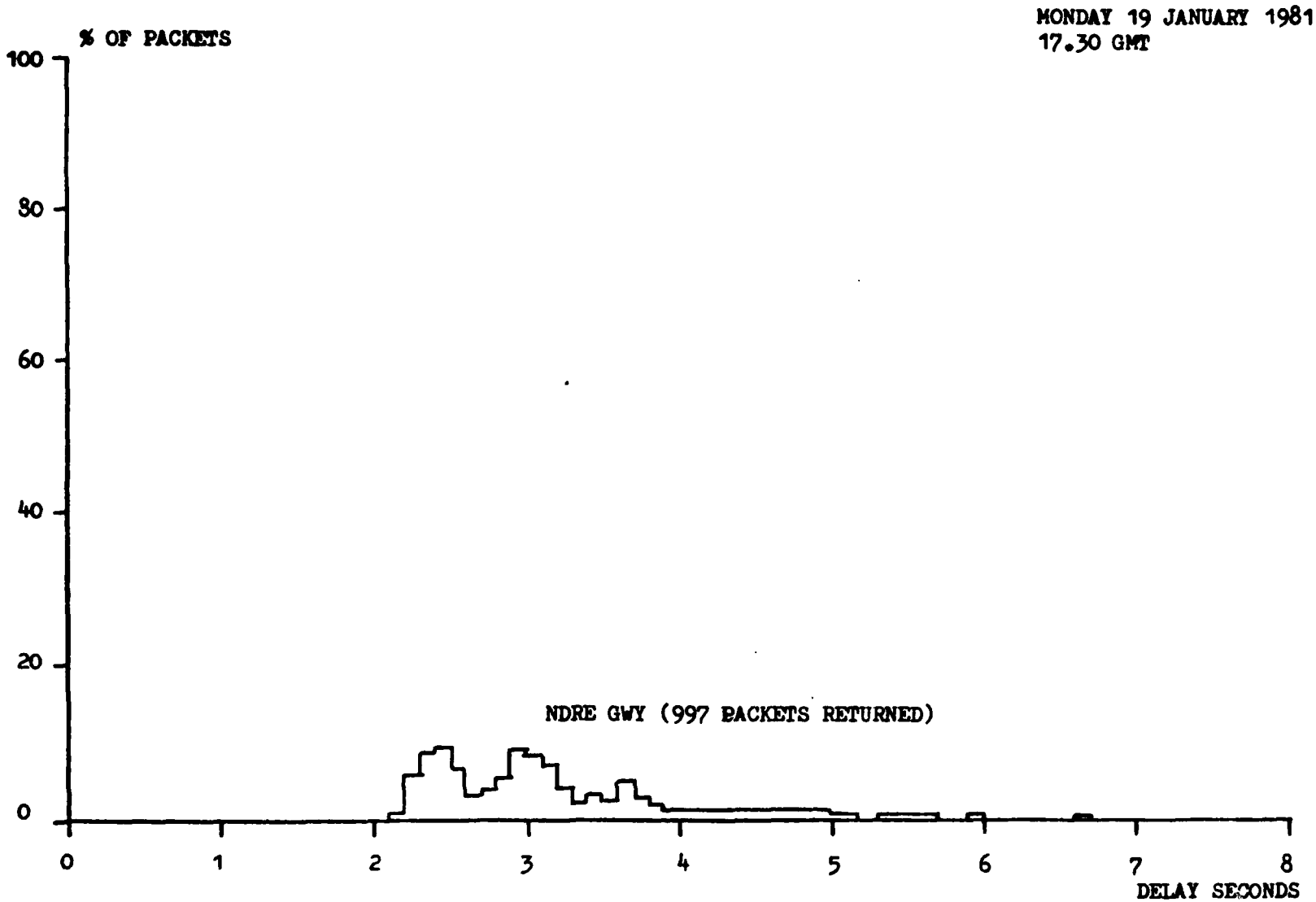


FIG 11. SINGLE ROUND TRIP DELAY MEASUREMENTS

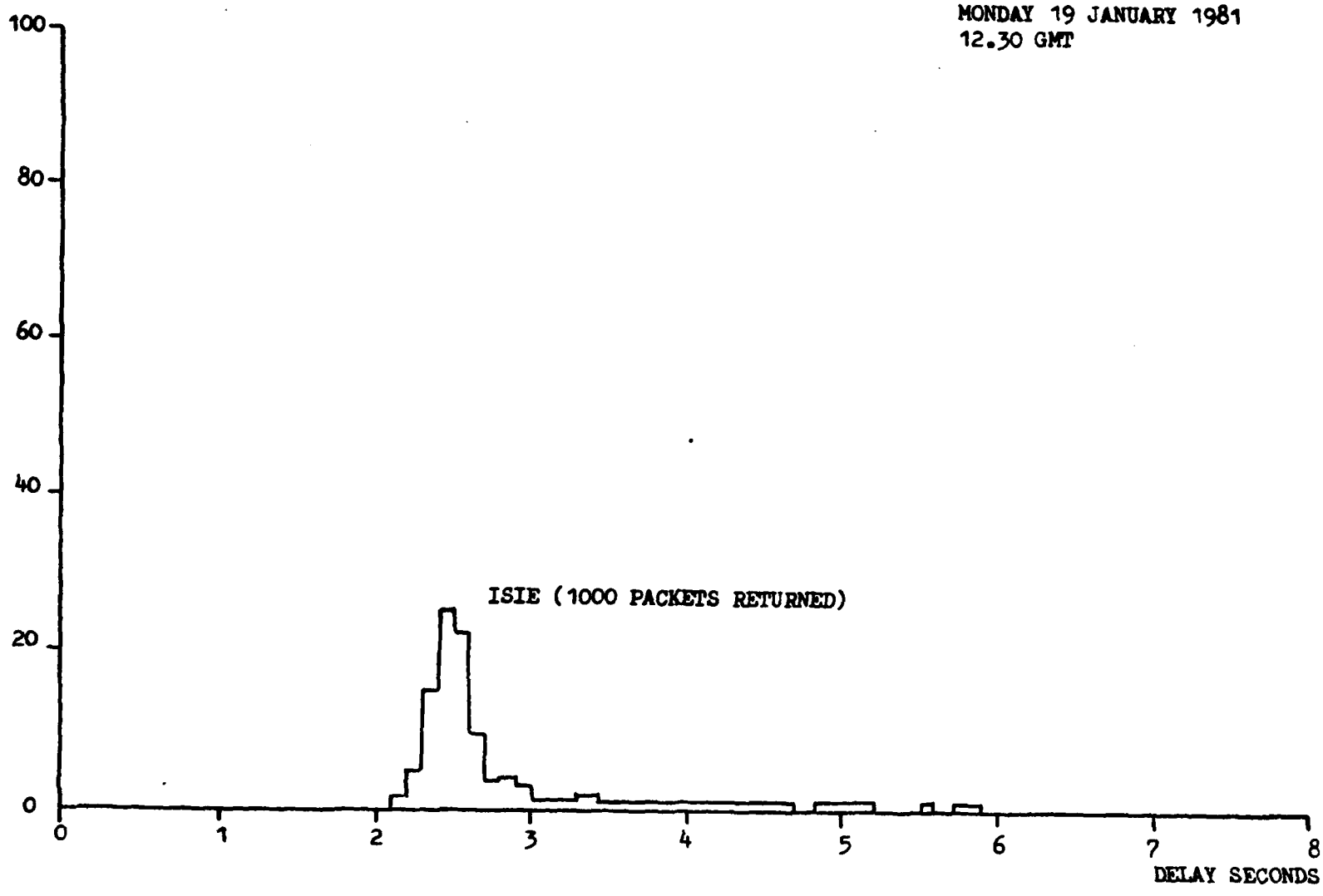


FIG 12. SINGLE ROUND TRIP DELAY MEASUREMENTS

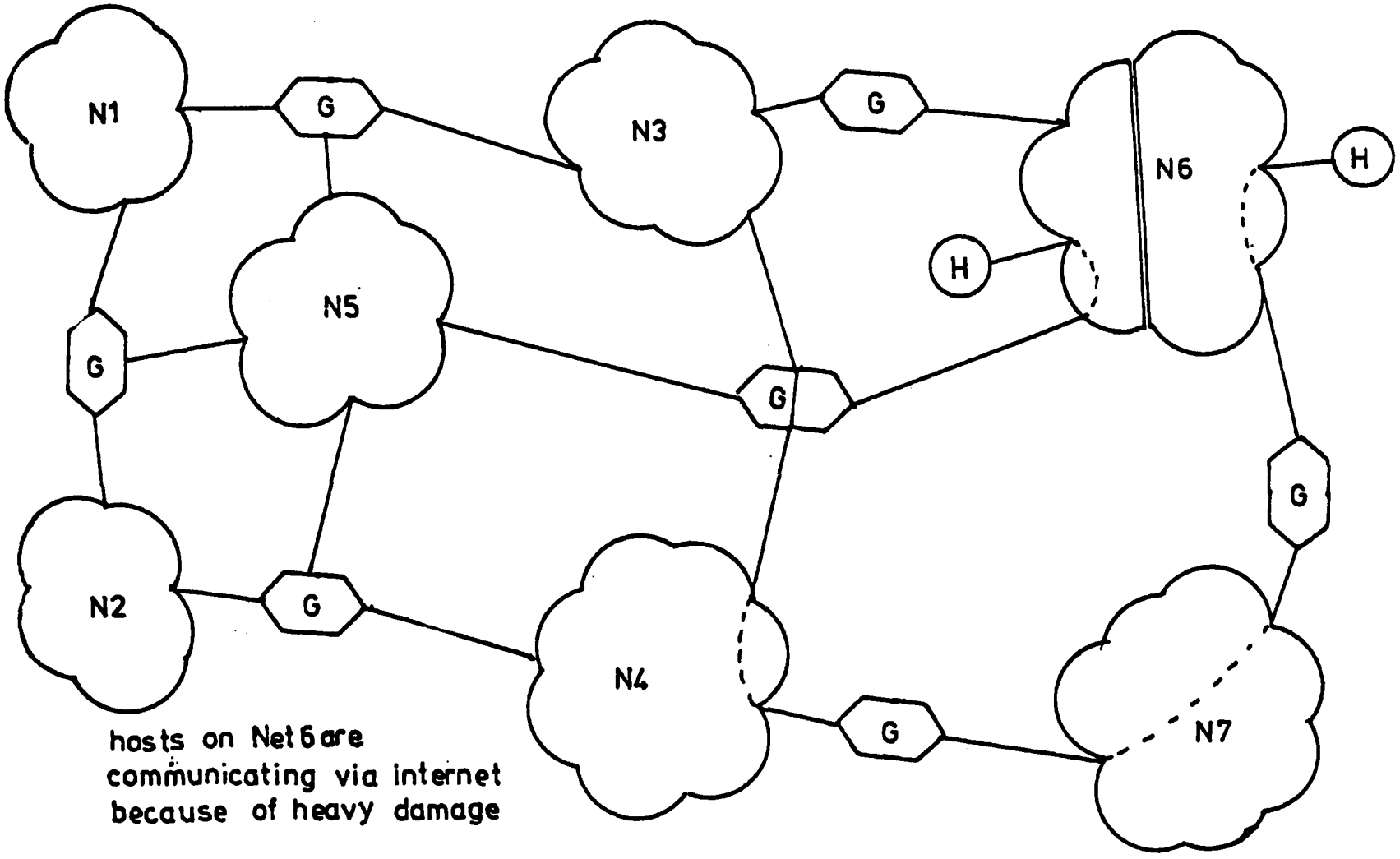


FIG.13 INTERNET SYSTEM CAN PROVIDE INCREASED COMMUNICATIONS SURVIVABILITY

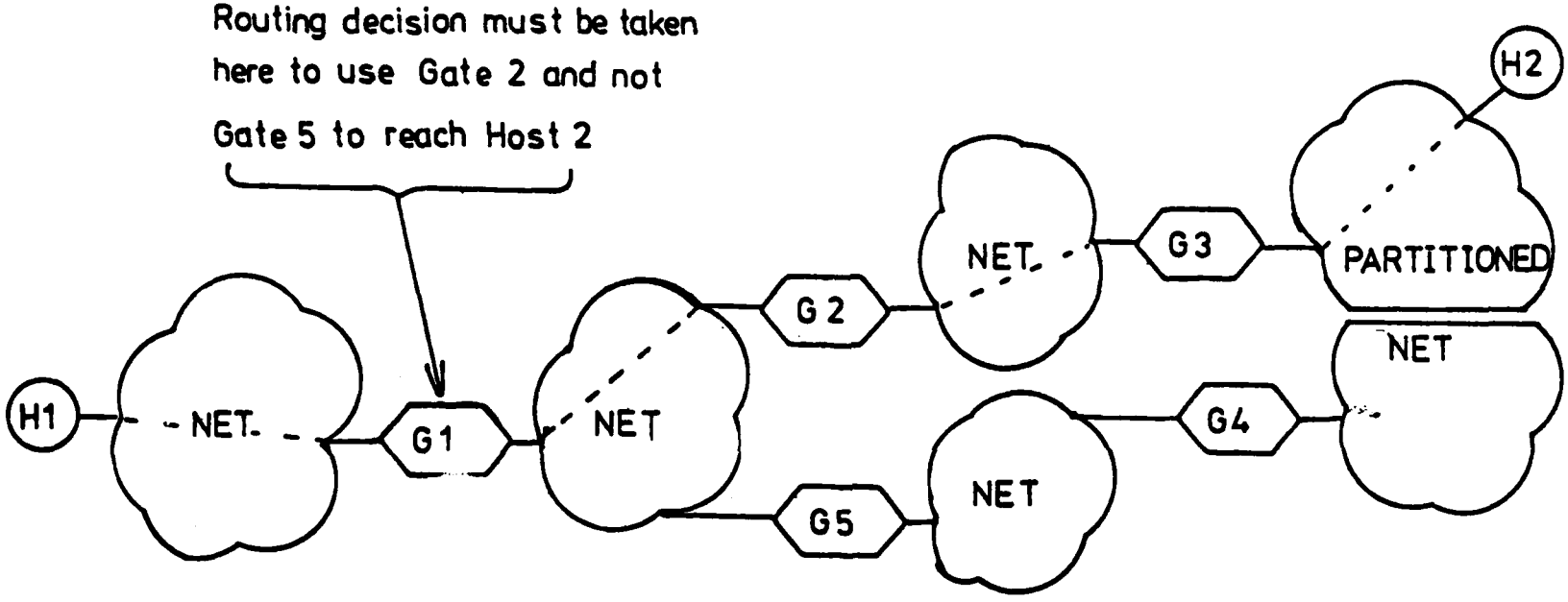


FIGURE14 ROUTING TO PARTITIONED NETWORKS

END

FILMED

2-83

DTIC