The Evolution of Packet Switching

LAWRENCE G. ROBERTS, MEMBER, IEEE

Invited Paper

***Abstract*—Over the past decade data communications has been revolutionized by a radically new technology called packet switching. In 1968 virtually all interactive data communication networks were circuit switched, the same as the telephone network. Circuit switching networks preallocate transmission bandwidth for an entire call or session. However, since interactive data traffic occurs in short bursts 90 percent or more of this bandwidth is wasted. Thus, as digital electronics became inexpensive enough, it became dramatically more cost-effective to completely redesign communications networks, introducing the concept of packet switching where the transmission bandwidth is dynamically allocated, permitting many users to share the same transmission line previously required for one user. Packet switching has been so successful, not only in improving the economics of data communications but in enhancing reliability and functional flexibility as well, that in 1978 virtually all new data networks being built throughout the world are based on packet switching. An open question at this time is how long will it take for voice communications to be revolutionized as well by packet switching technology. In order to better understand both the past and future evolution of this fast moving technology, this paper examines in detail the history and trends of packet switching.**

THERE HAVE ALWAYS been two fundamental and competing approaches to communications: pre-allocation and dynamic-allocation of transmission bandwidth. The telephone, telex, and TWX networks are circuit-switched systems, where a fixed bandwidth is preallocated for the duration of a call. Most radio usage also involves preallocation of the spectrum, either permanently or for single call. On the other hand, message, telegraph, and mail systems have historically operated by dynamically allocating bandwidths or space after a message is received, one link at a time, never attempting to schedule bandwidth over the whole source-to-destination path. Before the advent of computers, dynamic-allocation systems were necessarily limited to nonreal time communications, since many manual sorting and routing decisions were required along the path of each message. However, the rapid advances in computer technology over the last two decades have not only removed this limitation but have even made feasible dynamic-allocation communications systems that are superior to preallocation systems in connect time, reliability, economy and flexibility. This new communications technology, called "packet switching," divides the input flow of information into small segments, or packets, of data which move through the network in a manner similar to the handling of mail but at immensely higher speeds. Although the first packet-switching network was developed and tested less than ten years ago, packet systems already offer substantial economic and performance advantages over conventional systems. This has resulted in rapid worldwide acceptance of packet switching for low-speed interactive data communications networks, both public and private.

A question remains, however. Will dynamic-allocation techniques like packet switching generally replace circuit switching and other preallocation techniques for high-speed data and voice communication? The history of packet switching so far indicates that further applications are inevitable. The following examination of the primary technological and economic trade-offs involved in the growth of the packet switching communications industry should help to trace the development of the technology toward these further applications.

EARLY HISTORY

Packet switching technology was not really an invention, but a reapplication of the basic dynamic-allocation techniques used for over a century by the mail, telegraph, and torn paper tape switching systems. A packet switched network only allocates bandwidth when a block of data is ready to be sent, and only enough for that one block to travel over one network link at a time. Depending on the nature of the data traffic being transferred, the packet-switching approach is 3–100 times more efficient than preallocation techniques in reducing the wastage of available transmission bandwidth resources. To do this, packet systems require both processing power and buffer storage resources at each switch in the network for each packet sent. The resulting economic tradeoff is simple: if lines are cheap, use circuit switching; if computing is cheap, use packet switching. Although today this seems obvious, before packet switching had been demonstrated technically and proven economical, the tradeoff was never recognized, let along analyzed.

In the early 1960's, preallocation was so clearly the proven and accepted technique that no communications engineer ever seriously considered reverting to what was considered an obsolete technique, dynamic-allocation. Such techniques had been proven both uneconomic and unresponsive 20–80 years previously, so why reconsider them? The very fact that no great technological breakthrough was required to implement packet switching was another factor weighing against its acceptance by the engineering community.

What was required was the total reevaluation of the performance and economics of dynamic-allocation systems, and their application to an entirely different task. Thus, it remained for outsiders to the communications industry, computer professionals, to develop packet switching in response to a problem for which they needed a better answer: communicating data to and from computers.

THE PIONEERS

Rand

The first published description of what we now call packet switching was an 11-volume analysis, *On Distributed Communications*, prepared by Paul Baran of the Rand Corporation

Manuscript received May 11, 1978; revised May 22, 1978.
The author is with Telenet Communications Corporation, Washington, DC 20036.

0018-9219/78/1100-1307$00.75 © 1978 IEEE

in August 1964 [1]. This study was conducted for the Air Force, and it proposed a fully distributed packet switching system to provide for all military communications, data, and voice. The study also included a totally digital microwave system and integrated encryption capability. The Air Force's primary goal was to produce a totally survivable system that contained no critical central components. Not only was this goal achieved by Rand's proposed packet switching system, but even the economics projected were superior, for both voice and data transmissions. Unfortunately, the Air Force took no follow-up action, and the report sat largely ignored for many years until packet switching was rediscovered and applied by others.

ARPA I

Also in the 1962-1964 period, the Advanced Research Projects Agency (ARPA), under the direction of J. C. R. Licklider (currently at M.I.T.), sponsored and substantially furthered the development of time-sharing computer systems. One of Licklider's strong interests was to link these time-shared computers together through a widespread computer network. Although no actual work was done on the communication system at that time, the discussions and interest Licklider spawned had an important motivating impact on the initiators of the two first actual network projects: Donald Davies and me.

As previously indicated, the development of packet switching was primarily the result of identifying the need for a radically new communications system. Licklider's strong interest in and perception of the importance of the problem encouraged many people in the computer field to consider it seriously for the first time. It was in good part due to this influence that I decided, in November 1964, that computer networks were an important problem for which a new communications system was required [2]. Evidently Donald Davies of the National Physical Laboratory (NPL) in the United Kingdom had been seized by the same conviction, partially as a result of a seminar he sponsored in autumn 1965, which I attended with many M.I.T. Project MAC people. Thus, the interest in creating a new communications system grew out of the development of time-sharing and Licklider's special interest in the 1964-1965 period.

National Physical Laboratory

Almost immediately after the 1965 meeting, Donald Davies conceived of the details of a store-and-forward packet switching system, and in a June 1966 description of his proposal coined the term "packet" to describe the 128-byte blocks being moved around inside the network. Davies circulated his proposed network design throughout the U.K. in late 1965 and 1966. It was only after this distribution that he discovered Paul Baran's 1964 report.

The first open publication of the NPL proposal was in October 1967 at the A.C.M. Symposium in Gatlinburg, TN [3]. In nearly all respects, Davies' original proposal, developed in late 1965, was similar to the actual networks being built today. His cost analysis showed strong economic advantages for the packet approach, and by all rights, the proposal should have led quickly to a U.K. project. However, the communications world was hard to convince, and for several years, nothing happened in the U.K. on the development of a multinode packet switching network.

Donald Davies was able, however, to initiate a local network with a single packet switch at the NPL. By 1973 this local network was providing an important distribution service within the laboratory [4], [5]. This project, plus the strong conviction and continued effort by those at NPL (Davies, Barber, Scantlebury, Wilkinson, and Bartlett), did gradually have an effect on the U.K. and much of Europe.

ARPA II

In January 1967, I joined ARPA and assumed the management of the computer research programs under its sponsorship. ARPA was sponsoring computer research at leading universities and research labs in the U.S. These projects and their computers provided an ideal environment for a pilot network project; consequently, during 1967 the ARPANET was planned to link these computers together.

The plan was published in June 1967. The design consisted of a packet switching network, using minicomputers at each computer site as the packet switches and interfacing device, interconnected by leased lines. By coincidence, the first published document on the ARPANET was also presented at the A.C.M. Symposium in Gatlinburg, TN, in October 1967 [6] along with the NPL plan. The major differences between the designs were the porposed net line speeds, with NPL suggesting 1.5 Mbit/s lines. The resulting discussions were one factor leading to the ARPANET using 50-kbit/s lines, rather than the lower speed lines previously planned [7].

During 1968, a request for proposal was let for the ARPANET packet switching equipment and the operation of the network. The RFP was awarded to Bolt Beranek and Newman, Inc. (BBN) in Cambridge, MA, in January 1969. Significant aspects of the network's internal operation, such as routing, flow control, software design, and network control were developed by a BBN team consisting of Frank Heart, Robert Kahn, Severo Ornstein, William Crowther, and David Walden [8], [9], [10]. By December 1969. four nodes of the net had been installed and were operating effectively. The network was expanded rapidly thereafter to support 23-host computers by April 1971, 62 hosts by June 1974, and 111 hosts by March 1977.

The ARPANET utilized minicomputers at every node to be served by the network, interconnected in a fully distributed fashion by 50-kbit/s leased lines. Each minicomputer took blocks of data from the computers and terminals connected to it, subdivided them into 128 byte packets, and added a header specifying destination and source addresses; then, based on a dynamically updated routing table, the minicomputer sent the packet over whichever free line was currently the fastest route toward the destination. Upon receiving a packet, the next minicomputer would acknowledge it and repeat the routing process independently. Thus, one important characteristic of the ARPANET was its completely distributed, dynamic routing algorithm on a packet-by-packet basis, based on a continuous evaluation within the network of the least-delay paths, considering both line availability and queue lengths.

The technical and operational success of the ARPANET quickly demonstrated to a generally skeptical world that dynamic-allocation techniques—and packet switching in particular—could be organized to provide an efficient and highly responsive interactive data communications facility. Fears that packets would loop forever and that very large buffer pools would be required were quickly allayed. Since the ARPANET was a public project connecting many major universities and research institutions, the implementation and performance details were widely published [11], [12], [13],

[14], [15]. The work of Leonard Kleinrock and associates at UCLA on the theory and measurement of the ARPANET has been of particular importance in providing a firm theoretical and practical understanding of the performance of packet networks. (See "Principles and Lessons in Packet Communications" by L. Kleinrock, in this issue pp. 1320–1329.)

Packet switching was first demonstrated publicly at the first International Conference on Computer Communications (ICCC) in Washington, DC, in October 1972. Robert Kahn of BBN organized the demonstration. He installed a complete ARPANET node at the conference hotel, with about 40 active terminals permitting access to dozens of computers all over the U.S. This public demonstration was for many, if not most, of the ICCC attendees proof that packet switching really worked. It was difficult for many experienced professionals at that time to accept the fact that a collection of computers, wide-band circuits, and minicomputer switching nodes—pieces of equipment totaling well over a hundred—could all function together reliably, but the ARPANET demonstration lasted for three days and clearly displayed its reliable operation in public. The network provided ultra-reliable service to thousands of attendees during the entire length of the conference.

The widespread publicity the ARPANET demonstration earned contributed greatly to the task of introducing modern dynamic-allocation technology to a preallocation trained world. However, during the same period in the early 1970's many other dynamic-allocation techniques were being developed and tested in private networks throughout the world. Hopefully, the extensive publications on the ARPANET have not *oversold* the particular variety of packet switching used in this first major network experiment.

SITA

The Societe Internationale de Telecommunications Aeronautiques (SITA) provides telecommunications for the international air carriers. In 1969 SITA began updating its design by replacing the major nodes of its message switching network with High Level Network nodes interconnected with voice-grade lines—organized to act like a packet switching network. Incoming messages are subdivided into 240-byte packets and are stored and forwarded along predetermined routes to the destination. Prestored distributed tables provide for alternate routes in the event of line failures [16].

TYMNET

Also in 1969, a time sharing service bureau, Tymshare Corporation, started installing a network based on minicomputers to connect asynchronous timesharing terminals to its central computers. The network switches, which are interconnected by voice-grade lines, store and forward from node to node data characters for up to 20 calls packaged in 66-byte blocks. The data is repackaged at each node into new blocks for the next hop. Routing is not distributed, but is accomplished by a central supervisor on a call-by-call basis [17].

CYCLADES/CIGALE

In France the interest in packet switching networks grew quickly during the early 1970's. In 1973 the first hosts were connected to the CYCLADES network, which links several major computing centers throughout France. The name CYCLADES refers to both the communications subnet and the host computers. The communications subnetwork, called CIGALE, only moves disconnected packets and delivers them in whatever order they arrive without any knowledge or concept of messages, connections or flow control. Called a "datagram" packet facility, this concept has been widely promoted by Louis Pouzin, the designer and organizer of CYCLADES. Since a major part of the organization and control of the network is imbedded in the CYCLADES computers, the subnetwork, CIGALE, is not sufficient by itself. In fact, Pouzin himself speaks of the network as "including" portions of the host computers. The packet assembly and disassembly, sequence numbering, flow control, and virtual connection processing are all done by the host. The CYCLADES structure provides a good testbed for trying out various protocols, as was its intent; but it requires a more cooperative and coordinated set of hosts than is likely to exist in a public environment [18].

RCP

Another packet network experiment was started in France at about the same time by the French PTT Administration. This network, called RCP (Reseau a Commutation par Paquets), first became operational in 1974. By this time the French PTT had already decided to build the public packet network, TRANSPAC, and RCP was utilized primarily as testbed for TRANSPAC. The design of RCP, directed by Remi Despres, differed sharply from that of the other contemporary French network, CYCLADES. Despres' design was organized around the concept of virtual connections rather than datagrams. RCP's character as a prototype public network may have been a strong factor in this difference, since a virtual circuit service is more directly marketable, not requiring substantial modifications to customers' host computers. In any case, the RCP design pioneered the incorporation of individually flow-controlled virtual circuits into the basic packet switching network organization [19].

EIN

Organized in 1971 and originally known as the COST II Project and later as the European Informatics Network (EIN) is a multination-funded European research network. The project director is Derek Barber of NPL, one of the original investigators of packet switching in the U.K. Given freedom from the red tape of multinational funding, this project would have been one of the earliest pace-setters in packet networks in the world. As it happened, however, EIN was not operational until 1976 [20], [21].

Public Data Networks

The early packet networks were all private networks built to demonstrate the technology and to serve a restricted population of users. Besides those early networks already mentioned, which were the most public projects, many private corporations and service bureaus built their own private networks. Generally these private networks did not make provision for host computers at more than one location, and thus their organization usually developed into a star network.

All these networks were the result of a basic economic transition, which occurred in 1969 [22] when the cost of dynamic-allocation switching fell below that of transmission lines. This change made it economically advantageous to build a network of some kind rather than to continue to use direct lines or the circuit switched telephone network for interactive data communications. Universal regulatory conditions in all countries restricted "common carriage" to the government or government-approved carriers, and thereby led to the development of many private networks instead of a competitive market of public networks.

However, the extensive private network activity in the early 1970's encouraged some of these public carriers to make plans for building their own packet networks, although all public networks and plans for future networks were based on preallocation techniques until about 1973. Many plans to provide public data service arose; some were even under way, like the German EDS system; but all were based on circuit switching until that time. The shift in economics in the late 1960's that made packet switching more cost-effective, instigated more rapid change in communications technology than had ever before occurred.

The established carriers and PTT's took their time reacting to this new technology. The United Kingdom was the first country to announce a public packet network through the British Post Office's planned Experimental Packet Switched Service (EPSS) [23]. Donald Davies' 1966 briefings with the BPO on packet switching clearly played a strong role in the U.K.'s early commitment to this new technology.

In the United States the dominant carrier, American Telephone and Telegraph (AT&T), evidenced even less interest in packet switching than many of the PTT's in other countries. AT&T and its research organization, Bell Laboratories, have never to my knowledge published any research on packet switching. ARPA approached AT&T in the early 1970's to see if AT&T would be interested in taking over the ARPANET and offering a public packet switched service, but AT&T declined. However, the Federal Communications Commission (FCC), which regulates all communications carriers in the U.S., was in the process of opening up portions of the communications market to competition. Bolt Beranek and Newman, the primary contractor for the ARPANET, felt strongly that a public packet switched data communications was needed. The FCC's new policies encouraged competition, so BBN formed Telenet Communications Corporation in late 1972. In October 1973 Telenet filed its request with the FCC for approval to become a carrier and to construct a public packet switched network; six months later the FCC approved Telenet's request. (See "Legal, Commercial, and International Aspects of Packet Communications," by S. L. Mathison in this issue, pp. 1527-1539.)

In France in November 1973, the French PTT announced its plans to build TRANSPAC, a major domestic packet network patterned after RCP [24]. The next year, in October 1974, the Trans-Canada Telephone System announced DATAPAC, a public packet network in Canada [25]. Also during this period, the Nippon Telegraph and Telephone Corporation announced its plans to build a public packet switched data network in Japan [26].

Thus, only four years after the building of the first experimental networks, the concept of data communications networks began to move into the public arena. Still, the networks were only planned and had yet to be built; most PTT's and carriers adopted a wait and watch attitude toward these first public networks.

INTERNATIONAL STANDARDIZATION AND ACCEPTANCE

CCITT X.25

With five independent public packet networks under construction in the 1974-1975 period, there was strong incentive for the nations to agree on a standard user interface to the networks so that host computers would not have unique interfacing jobs in each country. Unlike most standards activities, where there is almost no incentive to compromise and agree, carriers in separate countries can only benefit from the adoption of a standard since it facilitates network interconnection and permits easier user attachment. To this end the parties concerned undertook a major effort, to agree on the host-network interface during 1975. The result was an agreed protocol, CCITT Recommendation X.25, adopted in March 1976.

The X.25 protocol provides for the interleaving of data blocks for up to 4095 virtual circuits (VC's) on a single full-duplex leased line interface to the network, including all procedures for call setup and disconnection. A significant feature of this interface, from the carriers' point of view, is the inclusion of independent flow control on each VC; the flow control enables the network (and the user) to protect itself from congestion and overflow under all circumstances without having to slow down or stop more than one call at a time. In networks like the ARPANET and CYCLADES which do not have this capability, the network must depend on the host (or other networks in interconnect cases) to insure that no user submits more data to the network than the network can handle or deliver. The only defense the network has without individual VC flow control is to shut off the entire host (or internet) interface. This, of course, can be disastrous to the other users communicating with the offending host or network.

Another critical aspect of X.25, not present in the proposals for a datagram interface, is that X.25 defines interface standards for both the host-to-network block transfer and the control of individual VC's. In datagram networks the VC interface is situated in the host computer; there can be, therefore, no network-enforced standard for labeling, sequencing, and flow controlling VC's. These networks are in the author's opinion, not salable as a public service since they must offer individual terminal interfaces, as well as host interfaces, to provide complete host-terminal communications; to sell these interfaces requires knowing how to interface to one VC as well as to a host.

The March 1976 agreement on X.25 and on virtual circuits as the agreed technique for public packet networks marked the beginning of the second phase of packet switching: large interconnected public service networks. In the two years since X.25 was adopted, many additional standards have been agreed on as well, all patterned around X.25. For example, X.28 has been adopted as the standard asynchronous terminal interface; X.29, a protocol used with X.25 to specify the packetizing rules for the terminal handler, will be the host control protocol. More recently X.75, the standard protocol for connecting international networks has been defined.

Public Data Network Services

Capitalizing on BBN's ARPANET experience, TELENET introduced the first public packet network service in August 1975. Initially TELENET consisted of seven multiply interconnected nodes. By April 1978 the network had grown to 187 network nodes which used 79 packet switches to provide 156 U.S. cities with local dial service to 180 host computers across the country, with interconnections to 14 other countries. Originally TELENET supported a virtual connection host interface similar to X.25. However, shortly after the specification was adopted, X.25 was introduced into TELENET as the preferred host interface protocol.

In early 1977 both EPSS in the U.K. and DATAPAC in Canada were declared operational. Also, in the U.S., TYMNET was approved as a carrier and began supplying public data services. EPSS, having been designed long before X.25 was specified, is not X.25 compatible, but the U.K. intends to provide X.25 based packet service within the next year.

DATAPAC was X.25 based from the start of commercial service since the development was held until X.25 was approved. Using X.25 lines, DATAPAC and TELENET were interconnected in early 1978. This connection demonstrated the ease of international network linking, once a common standard had been established.

In France, TRANSPAC is due to become operational later this year (1978); in Japan, the NTT packet network, DX-2, should become operational in 1978 or 1979. A semipublic network, EURONET, sponsored by nine European Common Market Countries, is due to become operational in late 1978 or 1979. Many other European countries, like Germany and Belgium, are making plans for public packet networks to start in 1979. These networks are all X.25 based and therefore should be similar and compatible.

Datagrams versus VC's

As part of the continuing evolution of packet switching, controversial issues are sure to arise. Even with universal adoption of X.25 and the virtual circuit approach by public networks throughout the world, there is currently a vocal group of users requesting a datagram standard. The two major benefits claimed for datagrams are reliability and efficiency for transaction-type applications.

Reliability: It is claimed that datagrams provide more reliable access to a host when two or more access lines are used, since any packet could take either route if a line were to fail. This reflects a true deficiency in X.25 as currently defined—the absence of a reconnect facility on the call request packet. If, when a call is initiated, a code number for the call is placed in the call request packet, the X.25 network (or host) can save the code number. If the line over which the call is placed fails, the network simply places a new call request, marked as a reconnect, over another line and supplies the original code number to insure reconnection to the correct VC. Since packets on each VC are sequence numbered, this reconnection can be accomplished with no data loss and usually just as quickly as rerouting of the packets in a datagram interface. If the network uses VC's internally, the same reconnect capability is used to insure against connection failures.

Cost: It is often assumed that datagrams would be cheaper for networks to provide than packets on VC's. However, the cost of memory and switching have fallen by a factor of 30 compared to transmission costs over the last nine years resulting in the overhead of datagram headers becoming a major cost factor. A datagram packet or end-to-end acknowledgment requires about 25 bytes of packet header in addition to the actual data (0–128 bytes) whereas only 8 bytes of overhead are required for similar packets on a virtual circuit. In the unique case of a single packet call, the overheads are the same. For all longer calls, datagram overhead adds 13–94 percent to the cost of all transmission costs, both long haul and local. Originally, this increase in transmission cost was more than offset by increased switch costs but with modern microprocessor switches very little of the increase is offset. Thus, with this radical shift in economics, datagram packets are now more expensive than VC packets.

CONTINUING TECHNOLOGICAL CHANGE

A decade ago computers had barely become inexpensive enough to make packet switching economically feasible; computers were still slow and small, forcing implementers to invent all sorts of techniques to save buffers and minimize CPU time. Computer technology has progressed to the point where microcomputer systems have now been especially designed for packet switching, and there is no shortage of memory or CPU power. This development has been partially responsible for the shift from datagrams to virtual connections and has also eliminated buffer allocation techniques (which cost transmission bandwidth to save memory). The modularity and computational power of today's microprocessors has made it economical and practical to provide protocol conversion from X.25 to any existing terminal protocol, polled or not.

As a result of these improvements, packet networks are rapidly becoming universal translators, connecting everything to everything else and supplying the speed, code, and protocol conversions wherever necessary. As this trend continues, it is almost certain that the techniques in use today will have to be continually changed to respond to the changing economics and usage patterns.

For example, one major change that will be required in the next few years is an increase in the backbone trunk speed from 56 kbit/s to 1.544 Mbit/s (the speed of "T1" digital trunks). Both Paul Baran and Donald Davies in their original papers anticipated the use of T1 trunks, but present traffic demand has not yet justified their use. As the traffic does justify T1 trunks, many aspects of network design will change by a corresponding order of magnitude. Packet networks have always incorporated a delay in the 100–200 ms range. This delay has so strongly affected both the system design and the choice of applications that it is hard to remember which decisions depended on this delay factor.

With T1 carrier trunks, the transit delay in the net will drop to around 10 ms plus propagation time requiring a complete reexamination of network topology and processor design issues. The number of outstanding packets on a 2500 mile trunk will increase from around 3 to 75, requiring extended numbering, and perhaps, new acknowledgment techniques. The user will be most strongly affected by a 10–30 ms net delay; his whole strategy of job organization may change.

Of course, there will be a significant price decrease accompanying this change. This, combined with the short delay, will make many new applications attractive; remote job entry (RJE) and bulk data transfer applications through public packet networks will probably be economically and technically feasible, even before T1 trunks are introduced; but if not before, certainly afterwards, when the packet price reflects the new trunk speed. Dynamic-allocation permits savings over pre-allocation by a factor of four in line costs for RJE, and by a factor of two for bulk data transfer. As the switch cost continues to fall far more rapidly than the line cost, dynamic-allocation techniques will be used for RJE, batch transfer and even voice applications.

FUTURE

Packet Satellite

One change which will clearly occur in packet networks in the next decade is the incorporation of broadcast satellite facilities. ARPA has sponsored extensive research into packet satellite techniques and, over the past few years, has tested these techniques between the U.S. and England. (See "General Purpose Packet Satellite Networks" by I. M. Jacobs, R. Binder, and E. Hoversten in this issue, pp. 1448-1467.) Fundamentally, a satellite provides a broadcast media which, if properly used, can provide considerable gains in the full statistical utilization of the satellite's capacity. Using ARPA's techniques, a single wideband channel (1.5 Mbit/s-60 Mbit/s) on a satellite provides an extremely economical way to interconnect high bandwidth nodes within a packet network.

With the current cost of ground stations ($150K-$300K), it appears to be marginally economic to install separate private ground stations at major nodes of a domestic packet network rather than to lease portions of commercial ground stations and trunk the data to the packet network nodes. However, either way, the cost of ground station facilities are such that the use of satellites only becomes economic compared to land lines when the aggregate data flow exceeds about 100 packets/s (100 kbit/s) to and from a node or city. Furthermore, satellite transmission has an inherent one-way delay of 270 ms; therefore, the packet traffic must logically be divided between two priority groups–interactive and batch. Only batch traffic can presently be considered for satellites, since the 270 ms delay is unacceptable for interactive applications, at least if any other options are available, even at a somewhat higher price. With current economics, the long-haul land line facilities only add about $0.50/hr to the price of interactive data calls, which is far too little a cost to encourage the acceptance of slower service. Therefore, interactive service will almost always require ground line facilities in addition to satellite facilities at all network nodes.

This introduces another factor that limits the potential satellite traffic: land lines can easily carry 10–25 percent batch traffic at a lower priority, using a dual queue, without any significant increase in cost. Further, if ground lines are required and satellite facilities are optional, the full cost for the satellite capability, must be compared with the incremental cost of simply expanding the land line facilities. All these factors considered, it is probable that satellites will be used by public data networks within the next five years for transmissions between major nodal points, but that ground facilities will be used exclusively for transmissions between smaller nodes.

Packet Radio

Since local distribution is by far the most expensive portion of a communications network, ground radio techniques are of considerable interest to the extent that they can replace wire for local distribution. Packet radio is another area where ARPA has been sponsoring research in applying dynamic-allocation techniques. The basic concept in packet radio is to share one wide bandwidth channel among many stations, each of which only transmits in short bursts when it has real data to send. (See "Advances in Packet Radio Technology" by R. E. Kahn *et al.*, in this issue, pp. 1468–1496.) This technique appears to be extremely promising for both fixed and mobile local distribution, once the cost of the transceivers has been reduced by, perhaps, a factor of ten. Considering the historical trend of the cost of electronics, this should take about five years; from that point onward packet radio should become increasingly competitive with wire, cable, and even light fibers for low to moderate volume local distribution requirements.

One important consequence would be the use of a simple packet radio system inside buildings to permit wireless communication for all sorts of devices. Clearly, as electronic devices multiply throughout the home and office, low-power packet radio would permit all these devices to communicate among themselves and with similar devices throughout the world via a master station tied into a public data network.

Voice

The economic advantage of dynamic-allocation over preallocation will soon become so fundamental and clear in all areas of communications, including voice, that it is not hard to project the same radical transition of technology will occur in voice communications as has occurred in data communications.

Digitized voice, no matter what the digitization rate, can be compressed by a factor of three or more by packet switching since in normal conversation each speaker is only speaking one third of the time. Since interactive data traffic typically can be compressed by a factor of 15, voice clearly benefits far less from packet switching than interactive data. This is the reason why packet switching was first applied to data communications. However, modern electronics is quickly eliminating any cost difference between packet switches and circuit switches, and thus packet switching can clearly provide a factor of three cost reduction in the transmission costs associated with switched voice service.

Probably there will be many proposals, and even systems built, using some form of dynamic-allocation other than packet switching during the period of transition. The most likely variant design would be a packetized voice system that does not utilize sum checks or flow control. Of course, this would be just a packet switch with those options disabled. If the similarity to present packet switching were not recognized, the packetized voice system might be built without providing these essential capabilities and would be useless for data traffic. However, the obvious solution would be an integrated packet switching network that provides both voice and data services.

On further consideration, it becomes apparent that the flow control feature of packet switching networks can provide a substantial cost reduction for voice systems. Flow control feedback, applied to the voice digitizers decreases their output rate when the network line becomes momentarily overloaded; as a result, peak channel capacity required by users can be significantly reduced.

In short, packet switching seems ideally suited to both voice and data transmissions. The transition to packet switching for the public data network has taken a decade, and still is not complete; many PTT's and carriers have not accepted its viability. Given the huge fixed investment in voice equipment in place today, the transition to voice switching may be considerably slower and more difficult. There is no way, however, to stop it from happening.

REFERENCES

[1] P. Baran *et al.*, "On distributed communications, vols. I-XI," RAND Corporation Research Documents, Aug. 1964.

[2] T. Marill and L. G. Roberts, "Toward a cooperative network of time-shared computers," *Proc. FJCC* pp. 425–431, 1966.

[3] D. W. Davies, K. A. Bartlett, R. A. Scantlebury, and P. T. Wilkinson, "A digital communications network for computers giving rapid response at remote terminals," ACM Symp. Operating Systems Problems, Oct. 1967.

[4] R. A. Scantlebury, P. T. Wilkinson, and K. A. Bartlett, "The de-

sign of a message switching centre for a digital communication network," in *Proc. IFIP Congress 1968, vol. 2–Hardware Applications,*" pp. 723-733.

[5] R. A. Scantlebury, "A model for the local area of a data communication network–Objectives and hardware organization," ACM Symp. Data Communication, Pine Mountain, Oct. 1969.

[6] L. G. Roberts, "Multiple computer networks and intercomputer communication," ACM Symp. Operating System Principles, Oct. 1967.

[7] L. G. Roberts and B. D. Wessler, "Computer network development to achieve resource sharing," *Proc. SJCC 1970*, pp. 543-549.

[8] F. E. Heart, R. E. Kahn, S. M. Ornstein, W. R. Crowther, and D. C. Walden, "The interface message processor for the ARPA computer network," in *AFIPS Conf. Proc. 36*, pp. 551-567, June 1970.

[9] R. E. Kahn and W. R. Crowther, "Flow control in a resource-sharing computer network," in Proc. Second ACM/IEEE Symp. Problems, Palo Alto, CA, pp. 108-116, Oct. 1971.

[10] S. M. Ornstein, F. E. Heart, W. R. Crowther, S. B. Russell, H. K. Rising, and A. Michel, "The terminal IMP for the ARPA computer network," in *AFIPS Conf. Proc. 40*, pp. 243-254, June 1972.

[11] H. Frank, R. E. Kahn, and L. Kleinrock, "Computer communications network design–Experience with theory and practice," in *AFIPS Conf. Proc.*, vol. 40, pp. 255-270, June 1972.

[12] L. Kleinrock, "Performance models and measurement of the ARPA computer networks," in *Proc. Int. Symp. Design and Application of Interactive Computer Systems*, Brunel University, Uxbridge, England, May 1972.

[13] L. Kleinrock and W. Naylor, "On measured behavior of the ARPA network," in *AFIPS Conf. Proc.*, vol. 43, NCC, Chicago, IL, pp. 767-780, May 1974.

[14] L. Kleinrock and H. Opderbeck, "Throughput in the ARPANET–Protocols and measurement," in *Proc. 4th Data Communications Symp.*, Quebec City, Canada, pp. 6-1–6-11, Oct. 1975.

[15] H. Frank and W. Chou, "Topological optimization of computer networks," *Proc. IEEE*, vol. 60, no. 11, pp. 1385-1397, Nov. 1972.

[16] G. J. Chretien, W. M. Konig, J. H. Rech, "The SITA network," in *Computer Communication Networks*, R. L. Grimsdale, F. F. Kuo, Eds. Noordhoff: NATO Advanced Study Institute Series, pp. 373-396, 1975.

[17] L. R. Tymes, "TYMNET–A terminal oriented communication network," *Proc. AFIPS, SJCC*, pp. 211-216, May 1971.

[18] L. Pouzin, "Presentation and major design aspects of the CYCLADES network," in *Third Data Communications Symp.*, Tampa, FL, pp. 80-85, Nov. 1973.

[19] R. F. Despres, "A packet network with graceful saturated operation," in *Proc. ICCC*, Washington, DC, pp. 345-351, Oct, 1972.

[20] D. L. A. Barber, "The European computer network project,"*Proc. ICCC*, Washington, DC, pp. 192-200, Oct. 1972.

[21] ——, "A European informatics network: Achievement and prospects," *Proc. ICCC*, Toronto, Canada, pp. 44-50, Aug. 1976.

[22] L. G. Roberts, "Data by the packet," *IEEE Spectrum*, Feb. 1974.

[23] R. C. Belton, J. R. Thomas, "The UKPO packet switching experiment," *ISS Munich*, 1974.

[24] A. Danet, R. Depres, A. LesRest, G. Pichon, S. Ritzenthaler, "The French public packet switching service: The TRANSPAC network," *Proc. ICCC*, Toronto, Canada, pp. 251-260, Aug. 1976.

[25] W. W. Clipsham, F. E. Glave, M. L. Narraway, "DATAPAC network overview," *Proc. ICCC*, Toronto, Canada, pp. 131-136, Aug. 1976.

[26] R. Nakamura, F. Ishino, M. Sasakoka, M. Nakamura, "Some design aspects of a public packet switched network," *Proc ICCC*, Toronto, Canada, pp. 317-322, Aug. 1976.