

HISTORY OF COMMUNICATIONS

EDITED BY MISCHA SCHWARTZ

AN EARLY HISTORY OF THE INTERNET

LEONARD KLEINROCK

INTRODUCTION

It is impossible to place the origins of the Internet in a single moment of time. One could argue that its roots lie in the earliest communications technologies of centuries and millennia past, or the beginnings of mathematics and logic, or even with the emergence of language itself. For each component of the massive infrastructure we call the Internet, there are technical (and social) precursors that run through our present and our histories. We may seek to explain, or assume away, whatever range of component technologies we like. It is equally possible to narrow Internet history down to specific technologies with which we are the most familiar.

There are also many individuals that may be said to have “predicted” the Internet. In 1908, Nikola Tesla foresaw [1] a technology that would allow “a business man in New York to dictate instructions, and have them instantly appear in type at his office in London or elsewhere” and would allow global access to “any picture, character, drawing, or print.” Thirty years later, H. G. Wells articulated [2] his idea of a “World Brain” as “a depot where knowledge and ideas are received, sorted, summarized, digested, clarified and compared.” These ideas were followed by a 1945 essay [3] by Vannevar Bush, predicting a machine with collective memory that he called the memex, with which “Wholly new forms of encyclopedias will appear, ready-made with a mesh of associative trails running through them, ready to be dropped into the memex and there amplified.”

These predictions, however, do not help us understand why the specific events, innovations, people, and circumstances that formed our Internet emerged when they did. Doing so is not possible from the scale of centuries or single individuals. This column’s focus is on the defining inventions and decisions that separate early technologies that were clearly not the Internet, from a wide range of recent inventions that may help characterize our Internet, but were also built within it. Thus, in this column we trace both the early history of the science and infrastructure that emerged as the ARPANET, and the trajectory of development it set for the even broader construct that we now call the Internet.

As one of many individuals who participated in the Internet’s early history, I

also offer a personal account of the same events, as an autobiographical element in this story. In doing so, I aim to further contextualize publications from the period — my primary source materials — with details from firsthand experience. This perspective may add to our depth of historical understanding, in which the extent of personal detail does not imply a greater importance to the events presented. In focusing on the work of individual researchers and developers, I rely on the various publications that followed the work of these individuals to link this story to the factual historical record we will follow. There are, of course, many important personal and institutional stories that have yet to be told. The University of California at Los Angeles (UCLA) is heavily mentioned in this column, as it was the site of so much foundational work. I view this period as a synergistic surge of technology, engineered by a magnificent group of researchers and developers amidst a defining period of challenge, creativity, invention, and impact.

BEFORE THE BEGINNING:
TWO THREADS THAT MEET

The Internet did not suddenly appear as the global infrastructure it is today, and neither did it form automatically out of earlier telecommunications. During the late 1950s and early 1960s, two independent threads were being woven. One was the research thread that eventually led to the packet switching networks of today’s Internet. This thread followed three possible paths to the technologies that eventually emerged; the researchers involved were, in chronological order, myself, Paul Baran, and Donald Davies. Below we explore these three paths, which were independently pursued in the quest to provide data networking theory, architecture, and implementation. The second thread was the creation and growth of the Advanced Research Projects Agency (ARPA), the institution that funded and deployed these technologies — a process that, as we will see, was by no means automatic. These two threads merged in the mid-1960s, creating the historical “break” that led to the ARPANET. Once these threads merged, the implementation and deployment phase began, bringing in other key contributors and successive

stages of development in Internet history. I present these threads and phases chronologically so we can revisit the history as it unfolded. One may find elaborations on this history in two earlier papers [4, 5]

THE RESEARCH THREAD

In January 1957¹ I began as a graduate student in electrical engineering at Massachusetts Institute of Technology (MIT). It was there that I worked with Claude Shannon, who inspired me to examine behavior as large numbers of elements (nodes, users, data) interacted; this led me to introduce the concept of distributed systems control and to include the study of “large” networks in my subsequent thesis proposal. In that MIT environment I was surrounded by many computers and realized that it would soon be necessary for them to communicate with each other. However, the existing circuit switching technology of telephony was woefully inadequate for supporting communication among these data sources. This was a fascinating and important challenge, and one that was relatively unexplored. So I decided to devote my Ph.D. research to solving this problem, and to develop the science and understanding of networks that could properly support data communications.

Circuit switching is problematic because data communications is bursty, that is, it is typically dominated by short bursts of activity with long periods of inactivity. I realized that any static assignment of network resources, as is the case with circuit switching, would be extremely wasteful of those resources, whereas dynamic assignment (I refer to this as “dynamic resource sharing” or “demand access”) would be highly efficient. This was an essential observation, and in 1959 it launched my research thread as I sought to design a new kind of network. Its architecture would use dynamic resource allocation to support the bursty nature of data communications, and eventually provide a structure for today’s packet-switched networks.

¹ Later that year on October 4, I experienced a widely shared feeling of surprise and embarrassment when the Soviet Union launched *Sputnik*, the first artificial Earth satellite. In response, President Eisenhower created ARPA on February 7, 1958 to regain and maintain U.S. technological leadership.

HISTORY OF COMMUNICATIONS

This concept of resource sharing was emerging at that time in a totally different context: that of timesharing of computer power. Timesharing was based on the same fundamental recognition that users generate bursty demands, and thus expensive computer resources were wasted when a computer was dedicated to a single user. To overcome this inefficiency, timesharing allocated the computer to multiple users simultaneously, recognizing that while one user was idle, others would likely be busy. This was an exquisite use of resource sharing. These ideas had roots in systems like SAGE [6] and in the MIT Compatible Time-Sharing System (CTSS [7]), developed in 1961 by Fernando Corbato (among the first timesharing systems to be implemented). The principles and advantages of timesharing were key to my realization that resource sharing of communication links in networks could provide for efficient data communications, much like the resource sharing of processors in timeshared systems was accomplishing.

In addition, there was already an example of a special-purpose data network that used resource sharing: the store-and-forward telegraph network. The challenge I faced was to create an appropriate model of general-purpose data communications networks, to solve for their behavior, and to develop an effective design methodology for such networks.

To do this, I sought to develop a model with dynamic resource sharing, incorporating the fact that data traffic was unpredictable as well as bursty. In order to clear up some misconceptions regarding what I and other investigators were doing in the field in the early days, I will devote some space in the following paragraphs to discuss the relationship between dynamic resource sharing and packet switching, where the latter is but one of many ways to realize the former. The basic structure I chose was that of a queue since it is a perfect resource sharing mechanism. A queue is dynamic, adaptive, and efficient, and does not wait for a message that is not there, but rather transmits a message already waiting in the queue. Moreover, the performance measures one considers in queueing theory are response time, throughput, efficiency, buffering, priorities, and so on, and these are just the quantities of interest in data networks. In the late 1950s, the published literature contained almost no work on networks of queues. However, a singular exception to this was the work by James Jackson, who published a classic paper [8] on open networks of queues.

As we see below, I was able to apply Jackson's result to represent the data networks of interest by making serious modifications to his model.

So the stage was set: There was a need to understand and design general-purpose data communication networks that could handle bursty data traffic, there was an emerging approach based on resource sharing in timeshared systems, there was an existing special-purpose network that suggested it could be done, and there was a body of queueing theory that looked promising.

As a result, I prepared and submitted my MIT Ph.D. thesis proposal [9] in May 1961, entitled "Information Flow in Large Communication Nets" in which I developed the first analysis of data networks. I chose a queueing theoretic model based on Jackson's model to characterize a data network as a network of communication channels whose purpose was to move data messages from their origin to their destination in a hop-by-hop fashion. Each channel was modeled as a resource serving a queue of data messages awaiting transmission; I discussed how "The nets under consideration consist of nodes, connected to each other by links. The nodes receive, sort, store, and transmit messages that enter and leave via the links..." My underlying model assumed that the stream of messages had randomly chosen lengths and, when applied to data networks, yielded a problem whose exact solution turned out to be hopelessly intractable. I altered the model and also introduced a critical mathematical assumption, the Independence Assumption,² which tamed the problem and allowed for an elegant solution. With this solution, I was able to solve for the many performance measures of these networks. For example, I showed that by scaling up the network traffic and bandwidth properly, one could reduce the system response time, increase the network efficiency, and increase the network throughput, all simultaneously [10].

In the course of examining data network performance, it became clear to me that it was important to explore the manner in which mean response time was affected when one introduced a priority queueing discipline on the traffic. I chose to understand this influence in the case of a single node first and then to apply the results to the general network case. This led to a publication in April

² Chapter 3 of my dissertation [16] elucidates this problem and the role of the Independence Assumption.

1962, which turned out to be the first paper [11] to introduce the concept of breaking messages into smaller fixed-length pieces (subsequently named "packets," as explained below). In it I provided a mathematically exact analysis of the mean response time, and showed the advantages to be gained by utilizing packet switching for this new network.³ Note that the fixed length packets I introduced did not match the randomly chosen lengths of the model, but fortunately, the key performance measure I solved for, the overall mean system response time, did not require that assumption, so the mathematical model properly reflected the behavior of fixed length packets as well.

I also developed optimal design procedures for determining the network capacity assignment, the topology, and the routing procedure. I introduced and evaluated distributed adaptive routing control procedures, noting that network/routing control is best handled by sharing control among all the nodes rather than relegating control to one or a small number of nodes. This distributes the control load (thereby not unduly loading any one node), introduces the ability to change routes on the fly dynamically (based on current load, connectivity, and destination address), enables the network to scale to a very large number of nodes, and dramatically improves the robustness of the network.

Whereas my focus was not principally on the engineering details of packet networks, I did address engineering details when I built a complete network simulation model and conducted extensive simulation experiments confirming the correctness of the theory. These experiments included detailed message blocks (with headers, origin and destination addresses, priority indicators, routing labels, etc), dynamic adaptive routing tables, priority queueing structures, traffic specifications, and more.⁴

Packetization was an integral part of a much broader body of knowledge that had to be developed to prove the case for data networks. Indeed, packetization alone was not the underlying technology

³ One of the important advantages of using packets turned out to be that short messages would not get "trapped" behind long messages; I was able to show this gain in response time exactly.

⁴ Access to my simulation notes can be found at http://ucla.worldcat.org/title/leonard-klein-rockscorrespondence-course-and-research-notes-1961-1972/oclc/263164964&referer=brief_results

HISTORY OF COMMUNICATIONS

that led to ARPANET design fundamentals. To be sure, packetization was and remains a core element of today's networking technology, but it is not identical to network efficiency. Rather, the fundamental gain lies in dynamic resource sharing. It is important to point out that there are many ways in which dynamic resource sharing can be accomplished, with packet switching being only one such method; other methods include polling [12], message switching [13], asynchronous time-division multiple access (ATDMA) [14], carrier sense multiple access with collision detection (CSMA/CD) [15], and others.

I completed and filed my Ph.D. dissertation [16] in December 1962, having created a mathematical theory of packet switching for dynamic resource sharing, thus providing the fundamental underpinnings for ARPANET technology. I showed that these networks were efficient, stable, scalable, robust, adaptive, and, most of all, feasible. Decades of important research on these topics have since taken place around the world.

By the time my dissertation was published as the first book [17] on computer networks in 1964, the idea of packetization itself was appearing more broadly. The next contributor to packet switching was Paul Baran of the RAND Corporation, who was busy working on military command and control systems during the early 1960s with the goal of using redundancy and digital technology to design a robust multilateral military communications network. He recognized the vulnerability of the telephone network due to its centralized architecture. In September 1962 he published a paper [18] on how "hot potato" adaptive alternate routing procedures and distributed principles could utilize a "standard message block," also to fall under the "packet" umbrella, which will be addressed below. His purpose was to create a network capable of functioning after a Soviet nuclear attack [19]. In August 1964 he produced a set of 11 important reports [20] reinforcing his prior description with simulations and elaborating on many details of the design. He, too, discovered the importance of going to digital networks and of the robustness provided by distributed routing. He attempted to get AT&T to implement the design, but failed to convince them (presumably due to their analog mindset). In 1965 RAND approached the Air Force to implement it, but they deferred to the DCA⁵; at this point, Baran decided not to pursue the implementation any further. Baran's

work was done independently of the work that I had done earlier at MIT and, in many ways, the results we achieved in addressing the problem of packet networks were complementary.

The third early contributor to packet switching was Donald Davies, of the National Physical Laboratory (NPL) in the United Kingdom. He began thinking about packet networks in 1965 and coined the term "packet" that year. In a privately circulated paper [21] dated June 1966, he described his design for a data network and used my earlier theory to calculate its performance. Davies lectured to a public audience in March 1967, recommending the use of his technology for the design of a public switched data network, and published an October 1967 paper [22] with his NPL group in which details of the design were first described in an open publication. This plan was for an NPL Data Communications Network, but the U.K. Department of Trade and Industry only authorized the implementation of one node. That node became operational in 1970. Further details of a full network design were described by the NPL team in 1968 [23, 24] and 1969 [25]; it is not clear where a multiple-node deployment by this team might have led, but it obviously had potential. This reluctance to support an NPL packet-switched network was reminiscent of the view taken by AT&T and DCA in not supporting an implementation of the RAND work.

The work of Baran and Davies focused on the engineering and architectural issues of the network design. My work emphasized and provided the mathematical underpinnings and supporting simulation experiments of the network analysis and design, including optimization as well as formulating the basic principles of packet networks that include dynamic resource sharing; this quantitatively showed that these networks were feasible. My trajectory was more fortunate as the ARPA thread rolled out and adopted my principles for their design of the ARPANET, and provided me the opportunity to participate in its implementation and deployment. Different trajectories were taken by Baran and then later by Davies, with Baran's unsuccessful attempts to get his ideas implemented and with Davies' frustration by the foot-dragging of the U.K. government. It was not enough to

put good ideas forward, but it was also necessary to prove that the concepts were quantitatively sound, and then to implement and deploy an operational network that would bring these ideas and designs to use.

THE ARPA THREAD

Let us step back chronologically and now pursue the second thread: the role of ARPA in defining the need for a data network, putting the management structure in place to enable its development, and providing the funding necessary for its implementation and deployment.

J. C. R. Licklider ("Lick") entered the story when he published his landmark 1960 paper [26] "Man-Computer Symbiosis." He defined the title as "an expected development in cooperative interaction between men and electronic computers." This work envisaged a system "to enable men and computers to cooperate in making decisions and controlling complex situations without inflexible dependence on predetermined programs"; he had seen such a flexible system in the aforementioned SAGE system. Once again, we find a forecast of what future telecommunications might provide — and Lick was perhaps the first to write at a time when viable ways to create that future were emerging. Although a visionary, Lick was not a networking technologist, so the challenge was to finally implement such ideas.

In May 1962 Lick and Welden Clark outlined their views on how networking computers could support social interaction, and provide networked access to programs and data [27]. This extended his earlier ideas of what he now referred to as a Galactic Network (in fact, he nicknamed his group of computer experts "The Intergalactic Network").

Lick was appointed as the first director of ARPA's newly formed Information Processing Techniques Office (IPTO) in October 1962. He quickly funded new research into advanced computer and networking technologies as well as areas that involved man-computer interaction and distributed systems.

By the end of 1962, Lick had articulated his grand vision for the Galactic Network, of which I was unaware, and I had laid out the mathematical theory of packet networks, of which Lick was also unaware. These ideas would soon intersect and reinforce each other in a series of key events between 1962 and 1969. I joined the UCLA faculty in 1963. Lick

⁵ Defense Communications Agency, which was renamed in 1991 to be today's (2010) Defense Information Systems Agency — DISA.

HISTORY OF COMMUNICATIONS

passed the directorship of IPTO to Ivan Sutherland, an MIT colleague of mine, in September 1964. In that role Sutherland wished to connect UCLA's three IBM mainframes in a three-node on-campus computer network, which would have been easy to accomplish with the means I had laid out in my Ph.D. dissertation. However, the UCLA network was never realized due to administrative discord. Nevertheless, the seeds for an ARPA-funded network had now been sown.

Early the next year (1965), Sutherland awarded Larry Roberts (another MIT colleague of mine who was quite familiar with my networking research) a contract to create a dialup 1200 b/s data connection across the United States. Later that year, Roberts accomplished this in collaboration with Thomas Marill, demonstrating that such a connection required a different, more sophisticated network than the telephone network offered [28].

Meanwhile, at ARPA, Sutherland recruited Robert Taylor to become associate director of IPTO in 1965. While there, Taylor also recognized the need for a network, this time specifically to connect ARPA research investigators to the few large expensive research computers across the country. This would allow them to share each other's hardware, software, and applications in a cost-effective fashion.⁶ Taylor then dropped into the office of the ARPA director, Charlie Herzfeld, to request funding for this nascent networking project. Herzfeld was a man of action who knew how to make a fast decision, and within 20 minutes he allocated \$1 million to Taylor as initial funding for the project. Taylor, who had since suc-

⁶ *This sharing of resources was the primary motivation for creating the ARPANET. Paul Baran developed a network design (described above) that would maintain communications — and specifically, Second Strike Capability — in the event of a nuclear attack by the USSR. His and my work served different aims. When ARPA began work on the ARPANET, my work was used for the reasons described herein. His application to military communications gave rise to the myth that the ARPANET was created to protect the United States in case of a nuclear attack. This is not to take away from Baran's accomplishments; indeed, by the time the ARPANET began in 1969, he had moved on to different projects, including the Institute for the Future (he stepped back into the ARPA foray in 1974–1975 to recommend that an early commercial version of the ARPANET be instituted beside the original research-driven network).*

ceeded Sutherland as IPTO director in August 1966, brought in Roberts as the IPTO chief scientist that December. Bringing Roberts in to manage the networking project turned out to be a critical hire as Roberts was to contribute at all levels to the coming success of data networking.

The research and ARPA threads had now merged, and the project would soon become the ARPANET.

These were critical steps in Internet history, for not even in the post-war United States did technological progress flow directly from ideas. In contrast to refusals from the private sector to fund the beginnings of the project, ARPA made available the will and funding of the U.S. government.⁷ ARPA's management and support fostered the early culture of shared, open research that was crucial to the success of the ARPANET program.

THE BEGINNING: THE ARPANET LAUNCH

The commitment to create the ARPANET was now in play. Roberts was empowered to develop the network concept based on Lick's vision, my theory, and Taylor's application.

There were basically two matters to be considered in this project. One was the issue of creating the switches and links underlying the network infrastructure, with the proper performance characteristics, including throughput, response time, buffering, loss, efficiency, scalability, topology, channel capacity, routing procedure, queueing discipline, reliability, robustness, and cost. The other was to create the appropriate protocols to be used by the attached (host) computers⁸ so that they could properly communicate with each other.

Shortly after his arrival, Roberts called a meeting of the ARPA Principal Investigators (PIs) in April 1967 at the

⁷ *In sharp contrast to ARPA's enthusiasm for networking, in the early 1960s, when I introduced the ideas of packet-switched networks to what was then the world's largest networking company, AT&T, I met with narrow-minded and failed thinking, and was summarily dismissed by them. They commented that packet switching would not work, and even if it did, they wanted nothing to do with it. Baran had a similar reaction from AT&T.*

⁸ *A major challenge for such a network was that it would connect computers with incompatible hardware and software.*

University of Michigan, where ARPANET planning was discussed in detail. It was there that the basic specifications for the underlying network were debated among us PIs. For example, Wesley Clark put forward the concept of using an unmanned minicomputer at each location to handle all of the switching and communications functions; it was to be called an Interface Message Processor (IMP). This would offload the networking functions from the host, greatly simplify the design by requiring only one interface to be written for each host to the standard IMP, and at the same time would decouple the network design from any specific host hardware and software. Another specification had to do with the measure of reliability of the planned network; this we specified by requiring that the topological design⁹ produce a "two-connected net," thus guaranteeing that no single failure would cause any non-failed portion of the network to lose connectivity.

Yet another requirement we introduced was for the network to provide an experience as if one were connected to a local timeshared computer even if that computer was sitting thousands of miles across the network; for this we specified that short messages should have response times no greater than 500 ms (the network design provided 200 ms at its inception). Moreover, since this was to start out as an experimental network, I insisted that appropriate measurement tools be included in the IMP software to allow for tracing of packets as they passed across the network, taking of snapshots of the IMP and host status at any time, artificial traffic generation, gathering and forwarding of statistics about the network, and a mechanism for controlling these measurements.

Following this April meeting, Roberts put together his outstanding plan for the ARPANET design and presented it as a paper [29] at a conference in Gatlinburg, Tennessee in October 1967. At this conference, Roger Scantlebury of the NPL also presented their aforementioned jointly published paper [22] describing a local network they were developing. It was during a conversation with Scantlebury at this meeting that Roberts first learned of the NPL work as well as some details of the work by Baran at RAND. The research by myself at MIT, by Baran at RAND,

⁹ *To assist with the topological design, Network Analysis Corporation (NAC), whose CEO was Howard Frank, was brought in as a contractor.*

HISTORY OF COMMUNICATIONS

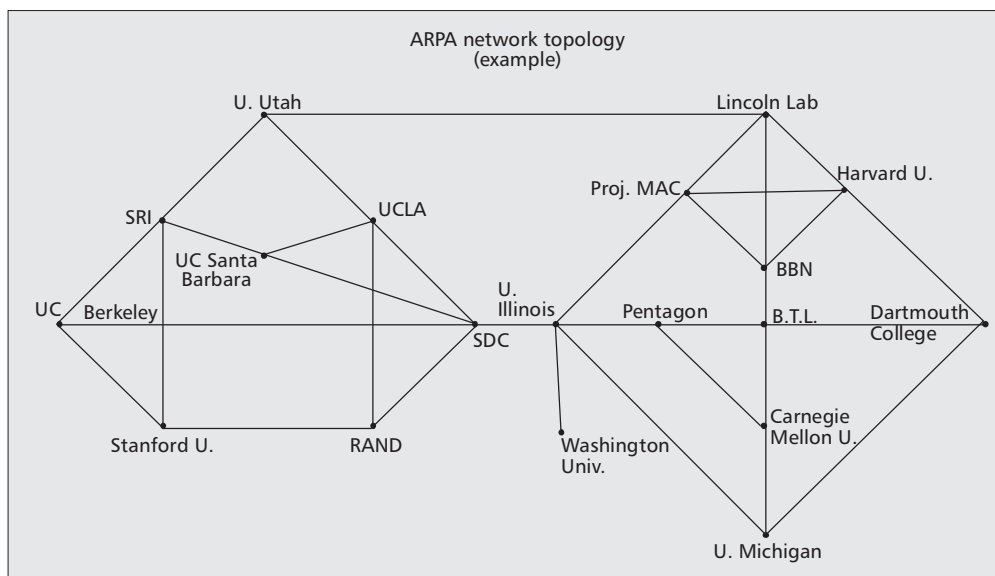


Figure 1. 19-node ARPANET as shown in the original RFQ.

and by Davies, Scantlebury, *et al.* at NPL had all proceeded independently, mostly without the researchers knowing about the others' work. There was, though, some cross-fertilization: Davies had used my analytical model for data networks in his work; as a result of discussions at this conference, Roberts adopted Davies' word "packet" for the small fixed length pieces I had suggested we break messages into, and which Baran referred to as "message blocks"; its fixed length was chosen to be 1024 bits for the ARPANET design (both Baran and Davies had suggested this same length); as a result of the discussion with Scantlebury, Roberts decided [30] to upgrade the backbone line speed from 9.6 kb/s to 50 kb/s for the ARPANET design.

Following these 1967 meetings, a sequence of drafts for the IMP specification was prepared.¹⁰ This culminated in March 1968 when Roberts and Barry Wessler produced the final version of the IMP specification, which they then discussed at an ARPA PI meeting later that month. On June 3, 1968, the ARPANET Program Plan [31] was formally submitted to ARPA by Roberts, and it was approved on June 21, 1968. The ARPANET procurement process was now officially underway.

By the end of July 1968, a Request

¹⁰ Among those involved in these first drafts were Frank Westervelt, Elmer Shapiro, Glen Culler, and myself.

for Quotation (RFQ) [32] for the network IMPs was mailed to 140 potential bidders. The 19-node example to be delivered by the contractor is shown in Fig. 1.

The handling of data streams specified that the hosts would communicate with other hosts by sending messages (of maximum length 8192 bits) to their attached IMPs, that these messages would be broken into packets (of maximum length 1024 bits each — thus, at most 8 packets per message) by the IMP, and that IMPs would communicate with each other using these packets. The movement of packets through the subnetwork of IMPs was to be controlled by a distributed dynamically updated routing algorithm based on network connectivity and loading as well as packet destination and priority. Errors in packet transmission between IMPs were managed by error detection and retransmission. Packets were to be reassembled into their original messages at the destination IMP before delivery to the destination host. The basic structure of this IMP specification contained contributions from a number of individuals, including my own research. Roberts had been well aware of my work since my time at MIT, where we were officemates, later stating,¹¹ "In order to plan to spend millions of dollars and stake my reputation, I needed to understand that it would work. Without Kleinrock's work of Networks and Queueing Theory, I could never have taken such a radical step." [33]

The RFQ resulted in 12 proposals being submitted in August 1968 (notably missing were IBM and AT&T). As these proposals were being evaluated at ARPA, Roberts awarded a research contract to me at UCLA in October to create the Network Measurement Center (NMC). The task of the NMC was to measure the behavior of the ARPANET by conducting experiments to determine its faults, performance, and outer limits (through the use of stress tests). I was fortunate to have a star team¹² of graduate student researchers, developers, and staff for this project; a number of these appear in continued roles later in this story. A week before Christmas 1968, Bolt, Beranek and Newman (BBN) won the competitive bid and was awarded the contract to develop the IMP-to-IMP subnetwork. The BBN team,¹³ supervised by Frank Heart, produced some remarkable accomplishments. This team had selected the Honeywell DDP-516 minicomputer with 12 kb of memory for the program to be the machine on which

¹¹ Roberts also goes on to say that my dissertation was "critical to my standing up to them and betting it would work."

¹² Key members of my UCLA team included a research team (Jerry Cole, Al Dobieski, Gary Fultz, Mario Gerla, Carl Hsu, Jack Zeigler), a software team (Vint Cerf, Steve Crocker, Gerard DeLoche, Charley Kline, Bill Naylor, Jon Postel), a hardware engineer (Mike Wingfield), and others.

HISTORY OF COMMUNICATIONS

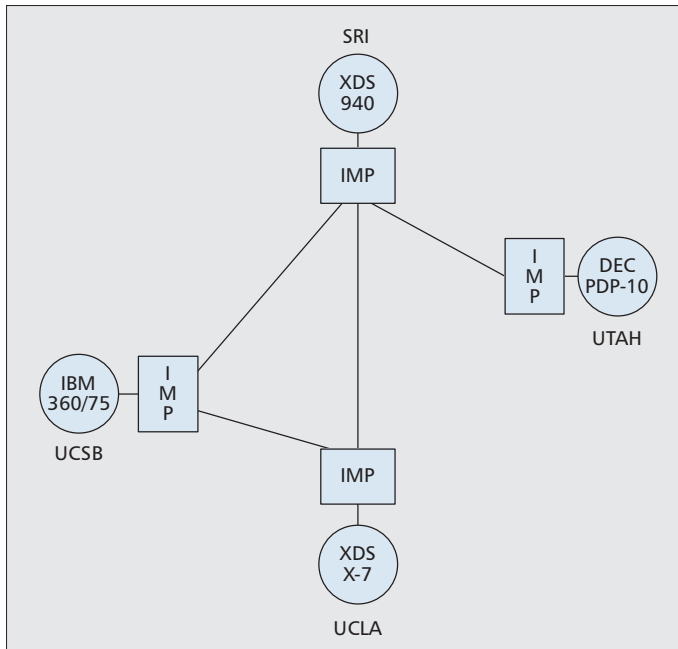


Figure 2. The initial four-node ARPANET (1969).

the IMP would be based; they were contracted to implement the IMP functions by modifying the hardware and software of the DDP-516, to connect these IMPs to long-haul 50 kb/s lines leased by Roberts from AT&T under the DoD Telpak tariff, and to deploy the subnetwork. The BBN team developed an elegant host-IMP design that met the ARPA specifications; this specification was written as BBN Report 1822 [34] by Robert Kahn, who was in charge of the system design at BBN (Kahn appears later in this story in some very significant roles, as we shall see below). One of the BBN team, Dave Walden, points out that he was most likely the first programmer on the Internet by virtue of having done code design for the IMP in their 1968 response to the RFQ. Whereas members of the BBN team were busy testing the IMP's ability to provide IMP-to-IMP data exchanges, testing the behavior of a network of IMPs was difficult to do in a laboratory environment; the true behavior was more properly tested in the deployed network with real traffic and with many nodes, which is exactly what the NMC was designed to do. Basically, BBN was given less than nine months to deliver the first IMP to

¹³ Key members of Heart's team included Ben Barker, Bernie Cosell, Will Crowther, Robert Kahn, Severo Omstein, Truett Thach, Dave Walden, and others.

UCLA by early September 1969. Their performance was outstanding. The first IMP at UCLA was to be followed by the second IMP in October to SRI, the third IMP in November to the University of California at Santa Barbara (UCSB), and the fourth IMP in December to the University of Utah. The initial network was to be that shown in Fig. 2.

These four sites were selected due to their ability to provide specialized network services and/or support. Specifically, UCLA (connecting an SDS Sigma-7 Host computer) would provide the NMC (under my supervision), SRI (connecting an SDS 940 host computer) would provide Doug Englebart's Human Intellect Augmentation System (with an early version of hypertext in his NLS system) as well as serve as the Network Information Center (under Elizabeth [Jake] Feinler's supervision), UCSB (connecting an IBM 360/75 host computer) would provide interactive graphics (under Glen Culler's and Burton Fried's supervision), and the University of Utah (connecting a DEC PDP-10 host computer) would provide advanced 3D graphics (under the supervision of Ivan Sutherland). The fact that Heart and his team at BBN succeeded in delivering this new technology with new applications and new users in an on-time, on-budget fashion was incredible.

But this contract to develop the underlying network was only the first of the two key tasks that were needed to

deploy a working packet-switched network. Recall that the other task was to create the appropriate protocols to be used by the attached (host) computers so that they could properly communicate with each other.

This second task was assigned to the four chosen ARPANET research sites to figure out on their own. Thus began another thread of innovative development that characterized the ARPANET culture. This thread actually begins in the summer of 1968 when Elmer Shapiro of SRI, in response to a request by ARPA, called a meeting of programmers from among those first sites that were to be connected into the ARPANET. Their main charge was to study and resolve the issues of host-to-host communication. Present at this meeting was one programmer from each of the first four sites to receive IMPs as follows: Steve Crocker (UCLA), Jeff Rulifson (SRI), Ron Stoughton (UCSB), and Steve Carr (University of Utah). This group, plus the many others who joined later, were soon to be named the Network Working Group (NWG) with Shapiro its first chairman.¹⁴ UCLA's Jon Postel served as the Request for Comments (RFC) editor (a role he held until his untimely death in 1998). They had no official charter against which to work, and so were afforded the unique opportunity to invent and create as needed. There was no sense of qualifying membership; all one had to do was to contribute and participate. Their focus moved to the creation of high level interactions and, eventually, to the notion of a layered set of protocols (transport services below a set of application-specific protocols). Basically, this was a highly resourceful, self-formed, collegial, loosely configured group of maverick graduate students who we (the ARPA PIs) had empowered to design and implement the protocols and software for the emerging network. They took on the challenge we ceded to them and

¹⁴ The names of some of the other key individuals who participated early on in the NWG include Bob Braden, Vint Cerf, Danny Cohen, Bill Duvall, Michel Elie, Jack Feinler, Jon Postel, and Joyce Reynolds.

¹⁵ It is remarkable how effective the RFCs, the NWG and the IETF have served the network community. In spite of the fact that they are loosely structured and involve large numbers of outspoken professionals, they have been able to move forward on a number of critical Internet issues.

HISTORY OF COMMUNICATIONS

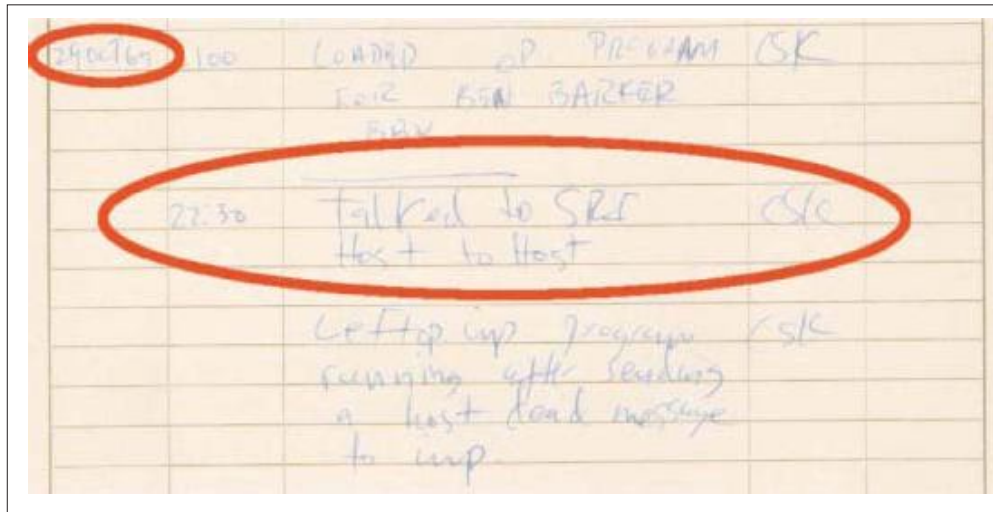


Figure 3. The entry in the original IMP log, which is the only record of the first message transmission on the Internet.

created an enduring NWG structure that later led to today’s Internet Engineering Task Force (IETF).¹⁵

Once the IMP-host specification was released by BBN in the spring of 1969, the NWG began to focus on the lower-level issues such as message formats. They decided to exchange ideas through a very informal set of notes they referred to as “Requests for Comments” (RFC). The first RFC [35], entitled “Host Protocol,” was written by Crocker in April 1969. Crocker became the second Chairman of the NWG early on.

We now had the two main ARPANET development efforts underway:

- A formal contract with BBN to create the IMP-IMP subnetwork
- An informal group of programmers (mostly graduate students) who were charged with developing the Host-to-Host Protocol

Things began to move rapidly at this point. The date of the first IMP delivery, scheduled to arrive to us at UCLA in early September 1969, was fast approaching. Meanwhile, at the NMC, we were busy collecting data so that we could predict performance of the network based on my earlier theory. For this, it was necessary to estimate the traffic loads that the host sites would present to the network. Roberts and I contacted a number of the early sites and asked them how much traffic they expected to generate and to which other sites. We also asked them how much traffic they would allow into their sites; to my surprise, many refused to allow any traffic from the network to use their

hosts. Their argument was that their hosts were already fully utilized serving their local customer base. Eventually they relented and provided their expected traffic loads. That traffic matrix was used in the July 1968 RFQ [32] and in a paper I published [36], thereby sealing their commitment.

On July 3, 1969, two months before the IMP was due to arrive, UCLA put out a press release [37] announcing the imminent deployment of the ARPANET. In that release I described what the network would look like, and what would be a typical application. I am quoted in the final paragraph as saying, “As of now, computer networks are still in their infancy, but as they grow up and become more sophisticated, we will probably see the spread of ‘computer utilities,’ which, like present electric and telephone utilities, will service individual homes and offices across the country.” It is gratifying to see that the “computer utilities” comment anticipated the emergence of web-based IP services, that the “electric and telephone utilities” comment anticipated the ability to plug in anywhere to an always on and “invisible” network, and that the “individual homes and offices” comment anticipated ubiquitous access. However, I did not foresee the powerful social networking side of the Internet and its rapidly growing impact on our society.

On Saturday, August 30, 1969, the first IMP arrived at UCLA. On September 2, the day after Labor Day, it was connected via a 15-foot cable to the UCLA host computer, our SDS Sigma-7 machine. This established the first

node of the fledgling network, as bits moved between the IMP and the Sigma-7. This is often regarded as a very significant moment in the Internet’s history.

In early October the second IMP was delivered by BBN to SRI in Menlo Park, California. The first high-speed link of what was to become the Internet was connected between those two IMPs at the “blazing” speed of 50 kb/s. Later in October, SRI connected their SDS 940 host computer to their IMP.

The ARPANET’s first host-to-host message was sent at 10:30 p.m. on October 29, 1969 when one of my programmers, Charley Kline, and I proceeded to “login” to the SRI host from the UCLA host. The procedure was for us to type in “log,” and the system at SRI was set up to be clever enough to fill out the rest of the command, adding “in,” thus creating the word “login.” Charley at our end and Bill Duvall at the SRI end each had a telephone headset so they could communicate by voice as the message was being transmitted. At the UCLA end, we typed in the “l” and asked SRI “did you get the l?”; “got the l” came the voice reply. We typed in the “o,” “did you get the o?,” and received “got the o.” UCLA then typed in the “g,” asked “did you get the g?,” at which point the system crashed! This was quite a beginning. So the very first message on the Internet was the prescient word “lo” (as in, “lo and behold!”). This, too, is regarded as a very significant moment in the Internet’s history.

The only record of this event is an

HISTORY OF COMMUNICATIONS

entry in our IMP log recording it as shown in Figure 3. Here we see that on October 29, 1969, at 10:30 pm, we at UCLA “Talked to SRI Host to Host.”

In November and December the IMPs and hosts at UCSB and the University of Utah were connected, respectively, thus completing the initial four-node network. Further IMP deliveries were halted until we had an opportunity to test this four-node network, and test it we did. Among other things, we were able to confirm with measurements some of our theoretical models of network delay and throughput as presented by Gerry Cole [38].

The ARPANET had now been launched. We now turn to the story of its rollout through its first decade.

THE FIRST DECADE: FOUR NODES AND THEN THE WORLD

By the time the first four nodes were deployed in December 1969, Roberts (who had succeeded Taylor in September to become the IPTO director) once again met with the NWG and urged them to extend their reach beyond what they had articulated in their first RFC [35], “Host Protocol.” This led them to develop a symmetric Host-to-Host Protocol, the first implementation of which was called the Network Control Program (NCP) and was described by Crocker in RFC 36 in March 1970 [39]. This protocol stack was to reside in the host machines themselves and included a hierarchy of layered protocols to implement more complex protocols. As NCP began deployment, the network users could begin to develop applications. The NCP was the first protocol stack to run on the ARPANET, later to be succeeded by TCP/IP. The trajectory of protocol stack development touched on below is another example of multiple possible paths that led the way from the ARPANET as it evolved into the Internet.

After the short evaluation period following the initial four-node deployment, a continual succession of IMPs and networks were then added to the ARPANET. In May 1970, at the AFIPS Spring Joint Computer Conference, a landmark session was devoted to the presentation of five papers [40] regarding the newly emerging ARPANET technology; these papers were packaged into a special ARPA pamphlet that was widely circulated in the community and spread information of the then-current technology that had been deployed.

(Two years later, in May 1972, another key session at the same conference was devoted to the presentation of five papers [41] that updated the ARPANET state of the art; this, too, was packaged into a second special ARPA pamphlet.) In mid-1970 the first cross-country link was added with a connection from UCLA to BBN, and by July the network contained 10 IMPs. The net grew to 15 IMPs by March 1971. In September 1971 BBN introduced a terminal interface processor (TIP) that conveniently would allow a terminal to connect directly to the ARPANET without the need to connect through an attached host. Later in the year, BBN slipped in a “minor” feature called electronic mail. Electronic mail had existed since the mid-1960s for standalone timeshared computer systems, but in late 1971 at BBN, Ray Tomlinson added a small patch to it that allowed the mail to pass between different computers attached to the ARPANET using an experimental file-sharing network program called CPYNET. Once he saw that it worked, he sent an email message to his group at BBN announcing this new capability, and so “The first use of network email announced its own existence.” [42]. This capability went out as a general TENEX release in early 1972. By July 1972, Roberts added a management utility to network email that allowed listing, selective reading, filing, forwarding, and replying to email messages. In less than a year email accounted for the majority of the network traffic. The network’s ability to extend communication between people was becoming evident, a nascent image of Lick’s vision.

Later that year, in October 1972, the first public demonstration of the ARPANET technology took place at the International Conference on Computer Communications (ICCC) in Washington, DC. Kahn, who by now had been hired into ARPA by Roberts, organized this large and very successful demonstration in which dozens of terminals in Washington accessed dozens of host computers throughout the United States in a continuously reliable fashion for the three-day duration of the conference.

The reaction of the computer manufacturers to this ARPANET phenomenon was to create proprietary network architectures based on their own brand of computers.¹⁶ The tele-

¹⁶ Among the proprietary networks were IBM’s SNA and DEC’s DECnet.

phone company continued to ignore it, but the open network that was the ARPANET thrived.

Soon, additional networks were added to the ARPANET, the earliest of which were those whose origins came out of work on wireless networking. Connecting the ARPANET with these different networks proved to be a feasible but not seamless interoperability issue, and it received a great deal of attention. The interconnection of networks was referred to as “internetworking” during the 1970s, a neologism from which the expanded ARPANET was eventually renamed as the Internet.

Let us briefly trace the work on wireless networking that led to these additional networks, which themselves forced attention on improving interoperability solutions. As pointed out above, these networks were based on wireless multi-access communications in which a shared channel is accessed by many users. By late 1970, Norm Abramson had developed AlohaNet [43] in Hawaii, a 9600 b/s packet radio net based on the novel “unslotted (pure) ALOHA” multi-access technique of random access. In this scheme (unsynchronized) terminals transmit their fixed length packets at any time over a shared channel at random times; if more than one transmission overlaps (i.e., collides), then destructive interference prevents any of the involved packets from succeeding. This tolerance of collisions was a departure from the more standard methods of wireline communications to control multi-access systems that used demand access methods (queueing, polling, etc., as mentioned earlier) and allowed only one transmission at a time (thus precluding such collisions). In 1973 Abramson calculated the capacity of the unslotted ALOHA system [44], which had a maximum efficiency of 18 percent, and in 1972 Roberts calculated the capacity of a synchronized version (i.e., slotted ALOHA) [45] whose capacity was doubled to 37 percent. However, these analyses ignored an essential issue with random access to shared channels: that they are fundamentally unstable, and some form of dynamic control was needed to stabilize them, for example, a backoff algorithm to control the way in which collided transmissions are retransmitted. This stability issue was first identified and addressed by Lam and myself [46, 47].

It is interesting to note that the ALOHA systems studies eventually led to an investigation of carrier sense multiple access (CSMA) as another wire-

HISTORY OF COMMUNICATIONS

less access method. CSMA itself led Robert Metcalfe to consider a variation called CSMA with collision detection (CSMA/CD), which was the basis for the original Ethernet development. Based on these concepts, Metcalfe and David Boggs implemented CSMA/CD on a coaxial cable network, which was up and running by November 1973. In sum, they created the Ethernet, which is today perhaps the world's most pervasive networking technology [48]. Ethernet is crucial to the story of NCP and TCP/IP, for researchers at Xerox PARC built on this technology in efforts to address the challenges of internetworking. Implemented in 1974 and published in 1975, the PARC Universal Packet (PUP) remained an internetwork architecture as late as 1979 [49]. PUP was one potential means through which to improve on NCP, although as we see below, that role was later taken on by TCP/IP. This is one of many stories that call out for more research into the histories and the individuals involved.

Let us now return to the story of the above-mentioned wireless technologies to help explain the motivation that led to TCP/IP (as different from that which motivated PUP). These technologies led to wireless networks that attached to the ARPANET, thereby exposing the nature of the problems of supporting connectivity among heterogeneous networks.

The first step was taken in December 1972, when an IMP in California used a satellite channel to connect to AlohaNet through an ALOHA host in Hawaii. Thus, the ARPANET, running the existing host-to-host Network Control Protocol, NCP, was now connected to a ground radio packet network, the AlohaNet. This was the first new network to connect to the ARPANET. AlohaNet had its own protocol and was working independent of ARPANET, yet a gateway provided internetwork connectivity between the two. In 1972 Roberts extended the ARPANET to Norway over a leased line that ARPA had already installed to receive seismic data and then extended it to London in the United Kingdom. This was the ARPANET's first international connection. In London Peter Kirstein then built a gateway to connect the ARPANET to a network built with another protocol between the U.K. universities. This was another case of different networks "internetworking," and as this function became an increasingly important focal point of ARPANET development, the network came to be known as the Internet to reflect this

growth. NCP was now handling the network-to-network interconnection of AlohaNet and the U.K. university network, both of which were attached to the ARPANET. The problems resulting from interconnected heterogeneous networks were becoming clear, and included the network-to-network protocol conversion needed between any (and every) pair of networks that were interconnected. It was clear that the combinatorial complexity of this pairwise protocol conversion would present considerable problems as the number of attached networks scaled up. TCP/IP was soon to emerge as the response chosen to address these problems.

At DARPA¹⁷ in early 1973, Kahn was the program manager responsible for, among other things, the ground packet radio network and the satellite packet radio network. He recognized the differences between the ARPANET running NCP, and these two radio networks. As a result, he set out to design a scalable end-to-end protocol that would allow dissimilar networks to communicate more easily. In the summer of 1973 Kahn discussed his approach for dealing with this internetwork complexity with Vint Cerf of Stanford who had considerable knowledge of NCP, since he had been a key member of the UCLA software group involved in the NCP design. Together, they drafted a detailed design of a new protocol, the Transmission Control Program (TCP). TCP was to take over the NCP's functions, but handle them in a more uniform manner: it would allow applications to run over an internetwork while hiding the differences between network protocols by using a uniform internetwork protocol. They distributed this design at a computer communications conference held at Sussex University in September 1973. (In October 1973 Roberts left IPTO to become CEO of TELENET, the first commercial packet switching network carrier.) By 1974, Cerf and Kahn fleshed out their design and published a definitive paper [50] on TCP. Underlying TCP was the key idea of an open network architecture that allowed packet networks of different types to interconnect with each other and for computers to exchange information end-to-end across these interconnected networks.

This contribution by Cerf and Kahn was a critical step in the development

¹⁷ ARPA was renamed DARPA in March 1972 when the word "Defense" was prepended.

of the Internet. In 1973–1974 DARPA commissioned three independent implementations of TCP: Cerf at Stanford University, Tomlinson at BBN, and Kirstein at University College London. In addition, David Clark of MIT worked on a compact version of TCP for the Xerox Alto personal workstation in the mid-1970s and later for the IBM PC desktop computer; David Reed, also of MIT, was working on internetworking among high-performance computers on LANs for the Laboratory for Computer Science Network (whose work was merged with the general TCP project in 1976). In August 1976 these implementations led to the first experimentation using TCP to connect two different networks: the packet radio network using Stanford's TCP implementation, and the ARPANET using BBN's TCP implementation. Following that, in 1977, Kahn implemented the satellite reservation protocol Roberts had designed, creating a second path from the ARPANET to the United Kingdom, sharing the capacity of a 64 kb/s Intelsat IV satellite broadcast channel among a number of ground stations in Europe and the East Coast of the United States. This Atlantic Packet Satellite Net (later to be called SATNET) was the ARPANET's second international connection. This was the second two-network TCP demonstration. Then a three-network demonstration of TCP was conducted on November 22, 1977, when the packet radio network, SATNET, and ARPANET were interconnected to allow an Internet transmission to take place between a mobile packet radio van at SRI and a USC/ISI host computer (both in California) via an intercontinental connection through University College London. This impressive feat was the first three-network TCP-based interconnection.

This first version of TCP only supported virtual circuits at the transport level (which is fine for applications that require reliable transmission). But it failed to support, among other things, real-time traffic such as packet voice where many aspects of the session flow were more properly handled by the application as opposed to the network. That is, real-time traffic called for support of an "unreliable" transport mechanism that would cope with missed packets, packets with errors, out-of-order packets, delayed packets, and so on. The use of unreliable transport support was already in use with NCP, prior to TCP; specifically, the early ARPANET IMP protocol allowed for

HISTORY OF COMMUNICATIONS

unreliable transport by use of what was called type 3 packets (also known as “raw” messages), which were introduced by Kahn in the BBN 1822 report. However, BBN was concerned that the uncontrolled use of these packets would degrade the network performance, so they regulated the use of type 3 packets to be on a limited, scheduled basis. In 1973–1974 Danny Cohen of USC/ISI implemented a Network Voice Protocol (NVP) [51] under ARPA support and requested BBN to allow him to use type 3 packets; with Kahn’s influence, BBN allowed this. Cohen’s real-time network voice experiments required the ability to cope with unreliable data transport. The early Version 1 design of TCP in 1974 did not support it, nor did Version 2 when it was implemented around 1977.

It was around this time that pressure for supporting unreliable transport in TCP came from Cohen, now joined by John Shoch and Reed, and with involvement from Crocker and Bob Braden. That is, they advocated modifying TCP such that type 3 packet functionality would be supported alongside reliable data transport. Cohen convinced Jon Postel of this, and Postel added a further concern, addressing layer violations, stating “We are screwing up in our design of internet protocols by violating the principle of layering. Specifically we are trying to use TCP to do two things: serve as a host level end-to-end protocol, and to serve as an Internet packaging and routing protocol. These two things should be provided in a layered and modular way. I suggest that a new distinct internetwork protocol is needed, and that TCP be used strictly as a host level end to-end-protocol.” [52] Postel then went on to describe how to break TCP into “two components: the hop-by-hop relaying of a message, and the end-to-end control of the conversation.” A robust internetworking solution was no easy task, and today’s TCP/IP was built with much experimentation on the ground laid by NCP.

Thus, there was a clear call to cleave TCP, splitting the function of network-layer connectivity, which involved addressing and forwarding, from its transport-layer end-to-end connection establishment, which also involved flow control, quality of service, retransmission, and more. TCP Version 3 (1978) introduced the split into two components, but it was only in TCP Version 4 (1980, with an update in 1981) that we see a stable protocol running that separated out the Internet Protocol (IP) from TCP (which now stood for Trans-

port Control Protocol) and was referred to as TCP/IP. This version has come to be known as IPv4. Along with the split into TCP and IP, the capability to support unreliable transport (i.e., type 3 packet functionality) was included. The formal name for this unreliable transport support was the User Datagram Protocol (UDP) [53].

In 1980 the U.S. Department of Defense (DoD) declared [54] the TCP/IP suite to be the standard for DoD. In January 1983, TCP/IP became the official standard [55] for the ARPANET; after a short grace period of a few months, no network was allowed to participate in the Internet if it did not comply with IPv4. Of course, Internet protocols never stop developing, and the 1998 upgrade to Version 6 dramatically extends the address space and introduces some significant security enhancements. It is still in the process of being deployed worldwide.

Meanwhile, as the 1970s rolled out, in addition to the ARPANET and TELENET, other packet networks were being designed across the globe in this period. Peter Kirstein, in his earlier paper [56] in this *IEEE Communications Magazine* History of Communications series, addresses much of the international work, especially the U.K. story (to which we refer the reader for more details). As a result of these national and international activities, an effort, spearheaded by Roberts, was put forth that resulted in the International Consultative Committee on Telephone and Telegraph (CCITT) Recommendation X.25. This agreed-upon protocol was based on virtual circuits — which was to be the CCITT’s own equivalent of TCP — and was adopted in 1976 [57]. During this period, the Network Measurement Center (NMC) at UCLA was deeply involved in measuring, testing, stressing, and studying the ARPANET. Bill Naylor and I published a summary of the tools used by the NMC as well as details of a weeklong measurement and evaluation of the results in 1974 [58]. In 1976 I published the first book that described the ARPANET technology, including its analytical modeling, design, architecture, deployment, and detailed measurements. A summary of the ARPANET principles and lessons learned appeared in a 1978 paper [59] after almost a full decade of experience with the use, experimentation, and measurement of packet networks; this paper was part of a special issue on packet communications which contains a number of key papers of that era [60]. One

of the first measurements we made was to determine the throughput from UCLA to UCSB in the initial four-node network shown in Figure 2; note that there are two paths between these two nodes. Whereas only one path was tagged as active in the routing tables at any one time, we found that both paths were carrying traffic at the same time since queued traffic continued to feed one of the paths when the other path was tagged. Among the more spectacular phenomena we uncovered were a series of lockups, degradations, and traps in the early ARPANET technology, most of which were unintentional and produced unpredicted side effects. These measurements and experiments were invaluable in identifying and correcting design issues for the early ARPANET, and in developing a philosophy about flow control that continues to inform us today. Moreover, it provided us, as researchers, a wealth of information for improving our theoretical models and analysis for more general networks. In July 1975 responsibility for the ARPANET was given to DCA. This terminated the systematic measurement, modeling, and stress testing that the UCLA NMC had performed for almost six years, and was never again restored for the Internet.¹⁸

It is outside the scope of this column to address Internet histories beyond those of its early period as the ARPANET. Likewise, I have not done justice to the untold stories that abound, but I hope to have convinced the reader that many people contributed to its success. This early history of the Internet, the first decade of design and deployment of the ARPANET, laid foundations on which today’s networks depend and continue to develop.

¹⁸ *The work of the NMC required a strong degree of cooperation from BBN since it was they who controlled any changes to the network code and architecture. At the NMC, each time we discovered a lockup, hardware problem, or other measured network problem, we alerted BBN so that they would take corrective action. Over time we developed an efficient working relationship with them, and errors were dealt with more expeditiously. It is worthwhile noting that the history of packet networks has met with institutional impediments to its progress, as have so many other technical advances over the course of history. In this case I have called out three with which I was personally involved: AT&T’s lack of interest in packet switching, the researchers’ reluctance to connect to the early network, and the above-mentioned negotiation with BBN.*

HISTORY OF COMMUNICATIONS

ACKNOWLEDGMENT

I want to thank Bradley Fidler, Mischa Schwartz, and the reviewers of this column for their helpful comments and suggestions.

REFERENCES (A PARTIAL LIST OF SELECTED REFERENCES)

- [1] W. W. Massie and C. R. Underhill, "The Future of the Wireless Art," *Wireless Telegraphy and Telephony*, 1908, pp. 67–71.
- [2] H. G. Wells, *World Brain*, Methuen & Co. Ltd., 1938.
- [3] J. Nye and P. Kahn, *From Memex to Hypertext — Vannevar Bush and the Mind's Machine*, Academic Press, Inc. 1991
- [4] L. Kleinrock, "History of the Internet and Its Flexible Future," *IEEE Wireless Commun.*, Feb. 2008, pp. 8–18.
- [5] B. Leiner et al., "The Past and Future History of the Internet," *Commun. ACM*, vol. 40, no. 2, Feb. 1997, pp. 102–08.
- [6] K. C. Redmond and T. M. Smith, *From Whirlwind to MITRE: The R&D Story of The SAGE Air Defense Computer*, The MIT Press, 2000.
- [7] F. J. Corbató, M. M. Daggett, and R. C. Daley, "An Experimental Time-Sharing System," *AFIPS Conf. Proc.*, vol. 21, 1962 SICC, pp. 335–44.
- [8] J. R. Jackson, "Networks of Waiting Lines," *Operations Research*, vol. 5, no. 4, Aug. 1957, pp. 518–21.
- [9] L. Kleinrock, "Information Flow in Large Communication Nets," Ph.D. thesis proposal, MIT, May 1961.
- [10] L. Kleinrock, "Creating a Mathematical Theory of Computer Networks," *INFORMS-Operations Research*, Jan.–Feb. 2002.
- [11] L. Kleinrock, "Information Flow in Large Communication Nets," RLE Quarterly Progress Report, MIT, Apr. 1962.
- [12] J. F. Kurose, M. Schwartz, and Y. Yemini, "Controlling Window Protocols for Time-Constrained Communication in a Multiple Access Environment," *Proc. 8th Symp. Data Commun.*, Oct. 3–6, 1983, pp. 75–84.
- [13] G. R. Andrews, "The Design of a Message Switching System: An Application and Evaluation of Modula," *IEEE Trans. Software Eng.*, vol. SE-5, No. 2, Mar. 1979.
- [14] W. W. Chu, "Demultiplexing Considerations for Statistical Multiplexors," *IEEE Trans. Commun.*, COM-20: 603–609, June 1972.
- [15] L. Kleinrock, *Queueing Systems, Volume II: Computer Applications*, Wiley Interscience, New York, 1976
- [16] L. Kleinrock, *Message Delay in Communication Nets with Storage*, Ph.D. dissertation, MIT, Cambridge, MA, 1962.
- [17] L. Kleinrock, *Communication Nets; Stochastic Message Flow and Delay*, McGraw-Hill, New York, 1964.
- [18] P. Baran, "On Distributed Communication Networks" Rand Paper P-2626, Sept. 1962.
- [19] D. Hochfelder and P. Baran, *An Oral History*, IEEE History Center, 24 Oct. 1999, http://www.ieeeahn.org/wiki/index.php/Oral-History:Paul_Baran#Institute_For_the_Future.2C_1967.2B
- [20] P. Baran et al., "On Distributed Communications," The RAND Corporation, Santa Monica, CA, Aug. 1964.
- [21] D. W. Davies, "Proposal for a Digital Communication Network," unpublished memo, June 1966, <http://www.archive.org/details/NationalPhysicalLaboratoryProposalForADigitalCommunicationNetwork>
- [22] D. W. Davies et al., "A Digital Communication Network For Computers Giving Rapid Response At Remote Terminals," *ACM Symp. Op. Sys. Principles*, Gatlinburg, TN, Oct. 1967.
- [23] D. W. Davies, "The Principles of a Data Communication Network for Computers and Remote Peripherals," *Proc. IFIP Hardware (Edinburgh1968) D1*.
- [24] R. A. Scantlebury, P. T., Wilkinson, and K. A. Bartlett, "The Design of a Message Switching Centre for A Digital Communication Network," *Proc. IFIP Hardware (Edinburgh 1968) D26*.
- [25] R. A. Scantlebury, "A Model for the Local Area of a Data Communication Network Objectives and Hardware Organization," *Proc. 1st ACM Symp., Problems in the Optimization of Data Communications Systems*, Pine Mountain, GA, Oct. 13–16, 1969, pp. 183–204.
- [26] J. C. R. Licklider, "Man-Computer Symbiosis," *IRE Trans. Human Factors in Electronics*, vol. HFE-1, Mar. 1960, pp. 4–11.
- [27] J. Licklider, and W. Clark, "On-Line Man-Computer Communication," *Spring Joint Comp. Conf.*, National Press, Palo Alto, CA, May 1962, vol. 21, pp. 113–28.
- [28] T. Marill and Lawrence Roberts, "Toward a Cooperative Network of Time-Shared Computers," *Fall AFIPS Conf.*, Oct. 1966.
- [29] L. Roberts, "Multiple Computer Networks and Intercomputer Communication," *ACM Symp. Operating System Principles*, Gatlinburg, TN, Oct. 1967.
- [30] See the fourth entry under "1967" in <http://www.packet.cc/>
- [31] L. Roberts, "Resource Sharing Computer Networks," June 3, 1968; http://ia310808.us.archive.org/load_djvu_applet.php?file=0/items/ResourceSharingComputerNetworks/Resource-SharingComputerNetworks.djvu
- [32] "Specifications of Interface Message Processors for the ARPA Computer Network," Department of the Army DBD/bj/OX 5-0494, July 29, 1968; http://www.cs.utexas.edu/users/chris/DIGITAL_ARCHIVE/ARPANET/RFQ-ARPA-IMP.pdf
- [33] L. G. Roberts, http://www.ziplink.net/users/roberts/SIGCOMM99/files/v3_document.htm
- [34] "Interface Message Processor: Specifications for the Interconnection of a Host and an IMP," BBN Report 1822; http://www.bitsavers.org/pdf/bbn/imp/BBN1822_Jan1976.pdf
- [35] S. Crocker, RFC 1, "Host Protocol"; <http://www.faqs.org/rfcs/rfc1.html>
- [36] L. Kleinrock, "Models for Computer Networks," Conference Record, *IEEE ICC*, Boulder, CO, June 1969, pp. 21-9–21-16.
- [37] T. Tugend, "UCLA to be first Station in Nationwide Computer Network," UCLA Press Release, 3 July 1969. <http://www.lk.cs.ucla.edu/LK/Bib/REPORT/press.html>.
- [38] G. C. Cole, "Computer Network Measurements; Techniques and Experiments," UCLA School of Engineering and Applied Science, Engineering Report UCLA-ENG-7165 Oct. 1971.
- [39] S. Crocker, RFC 36, "Protocol Notes"; <http://www.rfcarchive.org/getrfc.php?rfc=36>
- [40] Special session on the ARPANET in *Proc. Spring Joint Comp. Conf., AFIPS Conf. Proc.*, vol. 36, May, 1970, pp. 543–97.
- L. G. Roberts and B. D. Wessler, "Computer Network Development to Achieve Resource Sharing," pp. 543–49.
- F. Heart et al., "The Interface Message Processor for the ARPA Computer Network," pp. 551–67.
- L. Kleinrock, "Analytic and Simulation Methods in Computer Network Design," pp. 569–79.
- H. Frank, I. T. Frisch, and W. S. Chou, "Topological Considerations in the Design of the ARPA Network," pp. 581–87.
- S. Carr, S. Crocker, and V. Cerf, "Host-Host Communication Protocol in the ARPA Network," pp. 589–98.
- [41] Special session on the ARPANET in *Proc. Spring Joint Comp. Conf., AFIPS Conf. Proc.*, vol. 40, May, 1972:
- S. M. Ornstein et al., "The Terminal IMP for the ARPA Computer Network," pp. 243–54.
- H. Frank, R. Kahn, and L. Kleinrock, "Computer Communication Network — Design Experience with Theory and Practice," pp. 255–70.
- S. D. Crocker et al., "Function-Oriented Protocols for the ARPA Computer Network," pp. 271–79.
- R. H. Thomas, D. A. Henderson, "McROSS — A Multi-Computer Programming System," pp. 281–94.
- L. G. Roberts, "Extensions of Packet Communication Technology to a Hand Held Personal Terminal," pp. 295–98.
- [42] R. Tomlinson, "The First Network Email," <http://openmap.bbn.com/~tomlinso/ray/firste-mailframe.html>
- [43] N. Abramson, "The AlohaNet — Surfing for Wireless Data," *IEEE Commun. Mag.*, vol. 47, no. 12, Dec. 2009, pp. 21–25.
- [44] "The ALOHA System — Another Alternative for Computer Communication," *Proc. AFIPS Fall Joint Comp. Conf.*, vol. 37, 1970, pp. 281–85.
- [45] L. G. Roberts, "ALOHA Packet System with and without Slots and Capture," ARPA Network Info. Center, Stanford Research Inst., Menlo Park, CA, ASS Note 8 (NIC 11290), June 1972.
- [46] L. Kleinrock and S. Lam, "Packet Switching in a Slotted Satellite Channel," *AFIPS Conf. Proc.*, vol. 42, Nat'l. Comp. Conf., New York, June 1973, AFIPS Press, pp. 703–10.
- [47] S. Lam and L. Kleinrock, "Packet Switching in a Multiaccess Broadcast Channel: Dynamic Control Procedures," *IEEE Trans. Commun.*, vol. COM-23, Sept. 1975, pp. 891–904.
- [48] R. Metcalfe and D. Boggs "Ethernet: Distributed Packet-Switching for Local Computer Networks," *Commun. ACM*, vol. 19, no. 5, July 1976, pp. 395–404.
- [49] D. R. Boggs et al., "Pup: An Internetwork Architecture," *IEEE Trans. Commun.*, COM-28, no. 4, pp. 612–24, Apr., 1980, p. 26.
- [50] V. G. Cerf and R. E. Kahn, "A Protocol for Packet Network Interconnection," *IEEE Trans. Commun. Tech.*, vol. COM-22, V 5, May 1974, pp. 627–41.
- [51] D. Cohen, S. Casner, and J. W. Forgie, "A Network Voice Protocol NVP-II," ISI/RR-81-90, Apr. 1981.
- [52] J. Postel, "Comments on Internet Protocol and TCP," IEN 2, 1977.
- [53] J. Postel, "RFC 768 — User Datagram Protocol," 1980; <http://www.faqs.org/rfcs/rfc768.html>
- [54] G. P. Dinneen, "Host-to-Host Data Communications Protocols," Apr. 3, 1980; <http://www.potaroo.net/ietf/ien/ien152.txt>
- [55] J. Postel, "NCP/TCP Transition Plan," RFC 801, Nov. 1981; <http://www.rfceditor.org/rfc/rfc801.txt>
- [56] P. T. Kirstein, "The Early History of Packet Switching in the UK," *IEEE Commun. Mag.*, vol. 47, no. 2, Feb. 2009, pp. 18–26
- [57] T. Rybcznski, "Packet Switching (1975–1985): A Canadian Perspective," *IEEE Commun. Mag.*, vol. 47, no. 12, Dec. 2009, pp. 26–31.
- [58] L. Kleinrock and W. Naylor, "On Measured Behavior of the ARPA Network," *AFIPS Conf. Proc.*, vol. 43, National Comp. Conf., Chicago, IL, May 1974, AFIPS Press, pp. 767–80.
- [59] L. Kleinrock, "Principles and Lessons in Packet Communications," *Proc. IEEE*, Special Issue on Packet Communications, vol. 66, no. 11, Nov. 1978, pp. 1320–29.
- [60] *Proc. IEEE*, Special Issue on Packet Communications, vol. 66, no. 11, Nov. 1978.