45

THE STATUS AND PROMISE OF INTER-SITE COMPUTER COMMUNICATION[1]

M.W. Klopsch and S.G. Stafford

ABSTRACT

One major goal of the National Science Foundation's Long-Term Ecological Research (LTER) program is to facilitate collaborative research among sites. To achieve this goal, intersite communication of both data and documents will be required. Most of this information can be handled electronically. Data exchange via magnetic tapes, floppy disks, or telecommunications is available at all LTER sites; as communication speeds increase, nearly all data will be telecommunicated using error-checking protocols to improve reliability. Information currently communicated on paper could be sent as electronic mail; parts of collaborative manuscripts could be prepared on word processors and merged without extensive retyping or delays; electronically exchanged data summaries could easily be transformed for analysis or merged with other information. Shared databases could be established on either a central computer or an information service. Although the information revolution brings with it certain hazards, linking LTER sites into a telecommunication network is feasible and has great potential to enhance future ecological research.

INTRODUCTION

One major goal of the Long-Term Ecological Research (LTER) program is to facilitate collaborative research among

[1]FRL 1957, Forest Research Laboratory, Oregon State University, Corvallis. The mention of trade names or products does not constitute endorsement or recommendation for use.

sites. To achieve this goal, intersite communication of both data and documents is required. We surveyed LTER sites regarding their communications capabilities; the survey results suggest a considerable amount of unutilized communications potential.

Traditionally, most intersite communication has been by telephone conversations and paper documents--without the aid of electronic media. With the advent of the microprocessor, this situation has changed dramatically: most communications will soon be handled electronically at some point in their preparation. This fundamental change in operations provides the opportunity to speed and enhance intersite communication. In this paper we review current communication capabilities and explore some short-term and long-term possibilities.

THE CURRENT STATUS OF INTERSITE COMMUNICATIONS

Data exchange via magnetic tapes, floppy disks, or telecommunications is available at all LTER sites. Once a standard for tape interchange, a list of acceptable disk formats, and a knowledge of communication protocols have been established, data exchange between sites will be simply a matter of matching the method to the need.

Magnetic Tapes

Magnetic tape, a standard medium for mainframes and minicomputers, is useful for transferring large data sets and programs. Unless another format is arranged, the standard for magnetic tape interchange should be an "unlabeled," 9-track tape recorded in ASCII at 1600 bpi (bits per inch) with an ANSI format and a block size less than 5000 characters. Although the suggested format is somewhat bulky due to the large number of interblock spaces, it could be easily created and read by all LTER sites. The tape must be "unlabeled" because labeling installs a system-dependent password that makes reading the tape at another site nearly impossible. The block size must be small because some systems, such as our CDC CYBER computer, are unable to efficiently buffer larger blocks. Thorough documentation of tape formatting is required even with this standard. Our biggest problems with tape exchange have been caused by inadequate documentation of tape formats and the inadvertent exchange of "labeled" tapes.

Floppy Disks

Floppy disks are more commonly used with microcomputers, word processing systems, and to a lesser extent supermicrocomputers and minicomputers for medium-size data sets. Although disks are cheaper and simpler to mail than magnetic tapes, their use as a communication medium will primarily depend on the ease of moving the information to and from the disks at the origin and destination sites. The type of disk and format to use in disk exchanges should be determined by convenience. Although every LTER site can use 5¼-inch double-sided, double-density, MS-DOS formatted disks, many sites prefer other types of disks and formats (Table 1).

Table 1. Disk formats usable at LTER sites.

Site	Computer	Operating System	Disk Format	Format Acceptability
Andrews	Compupro	CP/M-80	8" SSSD	Preferred
Andrews	IBM PC	PC-DOS 2.1	5 ¼" DSDD	Preferred
Cedar Creek	IBM PC	PC-DOS 2.1	5 ¼" DSDD	Preferred
Cedar Creek	Xerox	CP/M-80	8" SSSD	Acceptable
Cedar Creek	Apple IIe	AppleDOS	5 ¼"	Possible
Cedar Creek	Apple IIe	ProDOS	5 ¼"	Possible
Coweeta	IBM PC	PC-DOS 2.1	5 ¼" DSDD	Preferred
Coweeta	MacIntosh	MacIntosh OS	3 ½"	Preferred
Coweeta	Apple IIe	AppleDOS	5 ¼"	Possible
Coweeta Hydro Lab	Apple IIe	AppleDOS	5 ¼"	Preferred
Illinois Rivers	IBM PC	PC-DOS 2.1	5 ¼" DSDD	Preferred
Illinois Rivers	Apple IIe	AppleDOS	5 ¼"	Possible
Jornada	IBM PC	PC-DOS 2.1	5 ¼" DSDD	Preferred
Jornada	MacIntosh	MacIntosh OS	3 ½"	Acceptable
Hobcaw	IBM PC	PC-DOS 2.1	5 ¼" DSDD	Acceptable
Hobcaw	Dec Pro 350	POS	5 ¼" RX50	Preferred
Konza Praire	IBM AT	PC-DOS 3.0	5 ¼" DSDD	Preferred
Konza Praire	IBM AT	PC-DOS 3.0	5 ¼" DSQD	Preferred
Konza Praire	CBM 8032	CP/M-80	5 ¼" SSSD	Acceptable
Niwot Ridge	IBM PC	PC-DOS 2.1	5 ¼" DSDD	Preferred
Niwot Ridge	RS Model 4	TRSDOS	5 ¼" DSDD	Preferred
Niwot Ridge	RS Model 16	XENIX	8" DSDD	Acceptable
Niwot Ridge	RS Model 3	TRSDOS	8" DSDD	Acceptable
Niwot Ridge	Apple IIe	AppleDOS	5 ¼"	Acceptable
Northern Lakes	IBM PC	PC-DOS 2.1	5 ¼" DSDD	Acceptable
Northern Lakes	Apple IIe	ProDOS	5 ¼"	Preferred
Northern Lakes	Apple IIe	P-System 1.2	5 ¼"	Acceptable
Northern Lakes	Apple IIe	AppleDOS	5 ¼"	Acceptable
Northern Lakes	MacIntosh	MacIntosh OS	3 ½"	Acceptable
Okefenokee	IBM PC	PC-DOS 2.1	5 ¼" DSDD	Preferred
Okefenokee	Viasyn Sys 10	CCPM	5 ¼" DSDD	Acceptable
Pawnee	IBM PC	PC-DOS 2.1	5 ¼" DSDD	Preferred
Kellogg	IBM PC	PC-DOS 2.1	5 ¼" DSDD	Preferred
Kellogg	VAX	VMS	8" files-II	Preferred

Telecommunications

Most communications among sites require rapidly transferring small amounts of data, for which telecommunications is best suited. In the future, as communication speeds increase, nearly all data will be handled by telecommunication.

Telecommunications can be either synchronous or asynchronous. Synchronous communication is the transmission of compressed blocks of data "in sync" with a clock signal. Standardized protocols are used to detect and correct transmission errors and to unblock the data. Synchronous modems are faster, but more expensive, than asynchronous modems and usually require specially conditioned phone lines. Only two sites reported the capability to telecommunicate synchronously with other sites. In contrast, asynchronous communication is based on the transmission of individual characters; each byte or character is transmitted by sending a starting bit, the data bits, sometimes a parity bit, and stop bits at a predetermined rate. Because asynchronous modems, software, and phone lines are cheaper than synchronous, most microcomputer communications are done asynchronously. Computers communicating asynchronously must use the same data configuration (*i.e.*, communication speed, data width, parity, stop bits, and half or full duplex). Most communication packages store these configurations along with the phone numbers to automatically configure the computer before dialing. The asynchronous configurations for LTER sites are reported in Table 2. Most asynchronous communication is conducted at a speed of 1200 baud (bits per second, about 120 characters), although some communication is still at 300 baud, and 2400-baud modems are becoming available. The most common 1200-baud configuration is eight data bits, no parity bit, and one stop bit, requiring ten bits to transmit a character.

The primary problems with telecommunications are comparatively high costs, relatively slow speed, and fairly high error rate. Since the reliability of telecommunications depends on the quality of the phone connection, it varies with location and time. The reliability problem can be overcome by using some form of error-checking protocol for important information. A standard protocol for intersite communication is desirable. The diversity of equipment among LTER sites requires that a protocol supported on a

Table 2. Intersite asynchronous communications capabilities.

Site	Telecommunications Contact Person	Phone	CS NET	Baud	Parity Data		Stop	Duplex	Error-Checking Protocols
H.J. Andrews	Mark Klopsch	(503) 757-4427	N	300/1200	7	E	1	Full	DATLINK, XMODEM, KERMIT
Cedar Creek	Robert Buck	(612) 376-9455	N	300/1200	7	E	1	Half	XMODEM
Coweeta	Polly Casale	(404) 542-2968	N	300/1200	7	E	1	Full	KERMIT (soon)
Coweeta Hydro Lab	Bryant Cunningham	(704) 524-2128	N	300/1200	8	N	1	Half	
Illinois Rivers	Frank Brookfield	(217) 333-6006	N	300/1200	8	N	1	Half	PLOT-10
Jornada	Walt Conley	(505) 646-2541	Y	300/1200	8	N	1	Half	Send twice and compare
Hobcaw	Robert McLaughlin	(803) 546-6219	N	300/1200	7	N	2	Full	
Konza Praire	John Briggs	(913) 532-6629	N	300/1200	8	N	2	Full	Hardcopy check
Niwot Ridge	Jim Halfpenny	(303) 492-6241	N	300/1200	8	N	1	Full	
Northern Lakes	Carl Bowser	(608) 262-8955	Y	300/1200	8	N	1	Full	Xon/Xoff
Okefenokee	Joe Schbauer	(404) 542-2968	N	300/1200	7	E	1	Full	KERMIT (soon)
Pawnee	Tom Kirchner	(303) 491-1986	Y	300/1200	7	E	1	Full	KERMIT
Kellogg	John Gorentz	(616) 671-5117	N	300/1200	7	O	1	Full	BTRANS, XMODEM

CS NET: Y = Currently Accessible, N = Not Currently Accessible
Parity: N = none, E = even, O = odd

wide variety of computers be selected. Only two asynchronous communication protocols currently deserve serious consideration: XMODEM and KERMIT. The most common, XMODEM, developed by Ward Christianson for his public domain MODEM7 package, is available on most microcomputers running CP/M-80, CP/M-86, MS-DOS, and ProDOS operating systems. Software providing this protocol include CROSSTALK, PC-TALKIII, MITE, and ASCOM (Helliwell, 1984). The KERMIT protocol, designed specifically for asynchronous communication between microcomputers and mainframes (DaCrus & Catchings, 1984a & b), is attractive because it is sufficiently simple to be easily implemented on most machines. Source code is available at cost from Columbia University for the IBM PC (PC-DOS), IBM 370 (VM/CMS), VAX-11 (VMS, UNIX), SUN (UNIX), PDP-11 (UNIX, RT-11, RSX, RSTS), 8080 (CP/M-80), 8086 (CP/M-86, MS-DOS), and Apple II (Apple DOS).

Information Services

Information services such as CompuServe and Source provide shared databases, electronic mail, bulletin boards, and forums on topics of general interest. The subscription, time-sharing, and telephone charges make the services relatively expensive ($10 to $50 per hour). Most of the features useful to the LTER sites could be approximated at a lower cost with a centrally located multi-user system.

THE POTENTIAL OF INTERSITE COMMUNICATIONS

Information currently communicated on paper could be sent more efficiently through an LTER electronic mail system, which could guarantee arrival in less than 24 hours for about the same cost as regular mail. For example, letters and memos, important news items, warnings, or requests for assistance could be routed electronically on the basis of names or keywords--perhaps to all principal investigators or to groups of people sharing a common interest. Parts of collaborative manuscripts could be prepared by their authors on word processors and merged without extensive retyping or delays. Electronically exchanged data summaries are easily transformed into spreadsheets or merged with other information into data files for future comparison and evaluation.

AT&T has recently been testing a new modem which takes advantage of the phone system's time compression multiplexing to permit full-duplex communication at 56K baud on ordinary phone lines (McDonnel, 1984). When the circuit-switched digital capacity (CSDC) system is universally available in late 1985, it should revolutionize telecommunications, replacing all other forms of communication, even for large data sets. The 20 cents it costs to telecommunicate a 2-page letter today should pay for the transfer of a 90-page manuscript a year from now. Facsimile transmission (the transmission of pictures and graphics), which also requires the transmission of large amounts of information, will be vastly enhanced. The CSDC system should make it possible to install reasonably efficient telecommunication bridges between local area networks (LANs) to achieve a functional regional, national, and international intercomputer communication network.

As the quality and speed of communications improve, certain graphics devices, specialty programs, and other network resources could be shared among sites. Because running a program on a distant time-sharing system is no more difficult than running it locally, many applications not requiring special display devices could be run remotely, reducing the need to replicate some expensive specialized equipment and software at each site. Shared databases, including lists of hardware and software with evaluations, data catalogs or dictionaries, and studies in progress, could be established on either a central computer or an information service like CompuServe.

Most LTER sites have access to a computer with the UNIX operating system. About 1500 large UNIX systems and many smaller computers are connected to UNIX communications program (UUCP) to form USENET (Emerson, 1983; Darwin, 1984). This program and associated utilities allow automatic file transfer between machines, electronic mail, and a subscription service for several hundred categories of news items. Each participant creates a subscription list identifying the news information categories desired. Because the UNIX system has provisions for scheduling unattended transfers, all USENET activities are accomplished with little or no supervision--a desirable feature.

However, as a network for LTER sites, USENET has several disadvantages. First, a message may pass through many computers before reaching its destination. For example, to get a message from Oregon State University to Emory University in Georgia, the message would have to pass through about ten machines. Though useful for passing on news and information, this is too slow for urgent messages and "rush" jobs. Second, USENET is so broadly based that, when it is used as a bulletin board, only general categories of news (such as physics, politics, and religion) are provided. Consequently, ecological researchers would have to wade through a morass of unnecessary information to find their topics of interest.

An alternative approach would be to set up a system among LTER sites in which the information categories could be specific to LTER needs. It might be possible to "piggyback" a smaller network of LTER sites onto the existing UNIX systems using the same software. An alternative would be to create an analogous network linking existing computer systems at each site. The disadvantage of developing an analogous network is that some software would have to be created to handle file forwarding and selection. A rudimentary system could be initiated using an asynchronous protocol with file-forwarding software located on only one central computer. In this star arrangement, users at each site would either manually or automatically contact the central computer node on a regular basis to exchange information. Later, the system could be enhanced with several interconnected central computers to reduce the telephone distances and to take advantage of faster or less expensive communications systems.

Some potential hazards accompany the proliferation of electronic publications and the use of electronic mail, data banks, and information services. First, users must remember that telecommunications is a tool, not a goal in itself, and must not be pursued to the exclusion of other projects ("network addiction"). Second, although freer exchange of information is a desirable goal, effective mechanisms must be developed to protect proprietary data and properly acknowledge individuals for their contributions. What constitutes a "publication," for example, may have to be redefined (Zientara, 1984). Finally, the possibility that we will all be overwhelmed by the sheer amount of information is a real danger. It is especially important for ecologists to learn the new tools for information access to avoid the narrowing of focus that would otherwise be almost inevitable.

The potential of efficient intersite communications is staggering. In the future, "idea exchanges" might be created in which hypotheses and supporting evidence could be displayed, comments or extra information added by other contributors, and an "electronic publication," with acknowledgments and authorship as appropriate for contributors, ultimately produced. The exchange could reduce research overlap and double as a registry for ideas to prevent piracy by associating an author and date with each entry. A new "gray" literature could be created if authors post papers for comment before publication. Increasingly, manuscripts for submission to journals or magazines will be requested in electronic form to speed galley preparation--a prelude to the electronic journals of the future, which promise a more timely exchange of data and ideas. Subscriptions to such journals could provide access to a database already cross-referenced by keywords and authors. The net result of these new procedures would be the rapid, organized dissemination of protected information, enhanced research opportunities, and improved journal quality.

Most LTER sites could be linked today through a simple communication network with existing equipment. Although current communication among sites might not seem to justify establishing a network, this is not a good indicator of future usage. In addition, implementing and standardizing communications within an LTER telecommunications network would enhance local communications at some sites. The LTER program is designed to provide leadership in ecological research. An LTER network would be a concrete example of this leadership and serve as a model or provide impetus for a global network of ecological researchers.

LITERATURE CITED

Darwin, I.F. 1984. The UNIX file. Microsystems 5(5): 32-38.

DaCrus, F. and B. Catchings. 1984a. Kermit: A file-transfer protocol for universities. Part 1: Design considerations and specifications. Byte 9(6): 225-278.

DaCrus, F. and B. Catchings. 1984b. Kermit: A file-transfer protocol for universities. Part 2: States and transitions, heuristic rules, and examples. Byte 9(7): 143-145, 400-403.

Emerson, S.L. 1983. USENET: A bulletin board for UNIX users. Byte 8(10): 219-236.

Helliwell, J. 1984. Guide to communications software. Computers & Electronics 22(8): 56-59, 86-89.

McDonnel, P. 1984. AT&T breaks the speed barrier. Computers & Electronics 22(9): 66-69, 100.

Zientara, M. 1984. Watch your words. Infoworld 6(32): 33-34.