```
01001100 01001100              01001100 01001100
01010100 01010100              01010100 01010100
01000101 01000101    LTER DATABITS    01000101 01000101
01010010 01010010              01010010 01010010
```

_____

Spring 1991   Data Management Newsletter of Long-term Ecological Research

This DATABITS features a wide array of articles and features with information on data acquisition, data manage-
ment systems, uses of LTER data and using GIS systems, plus our usual news from the sites. A new section, the
Technical Forum provides a look at some methods for implementing network-wide databases for selected datasets.

_____

_____

Contents

From the Sites

BNZ-- So as spring arrives and we dig ourselves out from a record
total snowfall (well over 12 feet and not only that but with an H2O
content WAY above normal), I sit here wondering how I'm supposed to
get everything done that I promised myself I would do before field
season (and our site visit)! I`ve also learned (what I already knew?)
that when you're also in charge of your own LAN don't kid yourself
into thinking that it won't consume a significant portion of your
time. We are in the process of purchasing Ingres RDBMS software to run
on the Sun system (a two node license for any of the 4 SPARCstations),
this will be used to develop our site LTER database. We can also use
the ARC/INFO RDBI-Ingres link to replace INFO with Ingres. I have
developed a series of forms, both paper and electronic for gathering
LTER database "meta-data". Now the trick is to get folks to use
them....I can see that I'm going to have to be inventive.

As some of you may know we are participating in the Inter-site Climate
Database proposal with the KBS site and others. Les Viereck and
Phyllis Adams have done an excellent job of keeping our climate
database in good shape (in spite of frustrating equipment problems)
and this project is the perfect opportunity to maximize utilization of
the data network-wide (lets hope it flies). --Mark Klingensmith,

Bonanza Creek LTER

CWT-- We are glad to announce that the renovations in the Institute of
Ecology is over, and that the Coweeta GIS and Data Management office
was given a very nice 350 sq. ft. room in the new space. Second, the
Image Analysis workstation has finally arrived. We have just finished
installing the scanner and the ERDAS software yesterday, and every-

thing seems to be working fine. We are planning to send mail to the
network describing this configuration as soon as we feel more
comfortable with it.   We think some of the other sites might be
interested. Finally, we are very excited to announce that the Coweeta
Site is now connected to internet. The bridge to the ` campus broad-
band was installed this morning. I couldn't wait to send you this
message.

I just would like to mention that we are NOT going to change our
current e-mail addresses in the next month or so.   All the work-
stations are going to be reconfigured to tighten up security holes,
and we might even change our host names. As soon as everything is
changed and running properly, I will contact Rudolf and let him know
about the new addresses. --Gil Calabria, Coweeta LTER

HFR-- Recently we've been working on the data management of GIS data:
developing conventions for GIS overlay names, and setting up databases
in Notebook (a PC text database) to document GIS map series and
individual overlays. The map series database includes such items as
name, number of rows & cols, cell size, and UTM coordinates; while the
overlay database contains such items as name, creation date, source
materials, procedures, updates, and labels. Initial experience with
this form of documentation has been encouraging. --Emery Boose,
Harvard Forest LTER

NIN-- The Marine Field Laboratory has still not yet been rebuilt since
it was destroyed in Hurricane Hugo. Blueprints are being refined but
construction has not yet begun. We are still temporarily housed in
several small buildings near the entrance to the property. The
Alphatronix Optical Drive serves beautifully for file system backups,
dumps and as a backup boot device. Evaluation of strategies to
implement SLIP from the Field Laboratory to the main campus continues.
Operation of the TOPS network of PC's, MAC's and the SUN has been
smooth. --Scott Chapall, North Inlet LTER

NTL-- We have been upgrading the LAN at the Limnology Laboratory which
is the administrative and data management headquarters for our site.
Installation of an Ethernet LAN which connects to the campus Ethernet
is progressing under the support from the 1990 supplement. The
Limnology Laboratory has been wired with twisted pair wiring to every
room to provide support of an Ethernet. Consultation with networking

experts at the central campus computing center led our planning committee to choose Novell Netware as the networking software because it can handle a heterogeneous environment and it is widely used and supported on the campus. The planning committee also decided to configure the file server on the Ethernet to have two hard drives and to operate with duplexing to protect against disk failure.

We are in the process of converting our fish and water chemistry databases into INGRES. One of our goals under the 1991 technical supplement is to acquire database management software for the local file server. We are interested in knowing which other sites are running Novell networks and if you have DBMS which operates in a heterogeneous environment. Also please contact us if your site uses INGRES. --Barbara Benson, North Temperate Lakes LTER

NWT-- We are currently in the process of developing a LAN. This work has proceeded more slowly than originally anticipated, primarily because we are coordinating our efforts with those of the other groups (Institute of Arctic and Alpine Research, Joint Facility for Regional Ecosystem Analysis, Cooperative Institute for Research in Environmental Sciences, Center for the Study of Earth from Space) within our building on the University of Colorado campus. All of the CULTER personnel within this building are affiliated with one or more of the other groups listed above. The networking of these groups is not a trivial matter since virtually all operating systems and platforms are represented. Moreover, the specific needs of the groups must be satisfied.

Nevertheless, we have made great progress during the past year and expect that our LAN will be operational within the next few months. The groups listed above have combined resources and negotiated with the university's Computing and Networking Services (CNS) for the purchase, installa- tion and maintenance of a Cisco router in the building. "Ethernet" cards or equivalent hardware have been purchased and installed in all CULTER microcomputers within the building. Thin-wire ethernet has been installed throughout the building and connections to the computers will be made during the next several weeks. We have acquired a "homegrown" software package from CNS that will provide immediate telnet and ftp capabilities. Our evaluation of the networking software package(s) that will meet our needs is ongoing. CNS has recommended that we postpone a decision until the key hardware components are in place and functional, however, since the market and available options are in constant flux.

CULTER is awaiting the delivery of the Sun SPARCstation II which should arrive the first week in April. Although this workstation will function primarily in a server capacity initially, it will eventually

provide access to all CULTER databases as well.

We are continuing to make improvements in computer-based communication and file transfers among RL1 (the building where most CULTER offices are located; approximately 1 mile east of the main campus), Ramely Hall (the building where several CULTER investigators and graduate students have their offices; on main campus), and the Mountain Research Station (our remote field station and laboratory; approximately 25 miles northwest of Boulder). Although electronic communication and file transfer capabilities among these three sites have existed for some time, we are increasing the number of micro-computers with these capabilities as well as enhancing the speed and efficiency of transfers. The Mountain Research Station was recently donated a new telephone system and the modems being replaced by ethernet connections are being moved to MRS so that more of the microcomputers "up there" can talk to us "down here". --Rick Ingersoll, Niwot Ridge LTER

VCR-- We are trying to recover from a major loss: Bill Odum, our recently installed PI, died in early April as a result of an untreatable liver ailment. Bill meant a great deal (both personally and professionally) to the researchers of the VCR/LTER and will be much missed.

We are in the middle of "Project Description" season, as PI's and

students prepare documentation for their summer field season activities and progress reports. The project descriptions are then stored in a DBASE IV-based database for inclusion in dataset descriptions and annual reports. We have also been exploring collaborative opportunities with elements of the Computer Science Department at the University of Virginia. Anita Jones, chairman of the department, James French and six students attended a meeting to discuss technical areas of data management that needed to be advanced, possibly as thesis projects for master's students. -- John Porter, Virginia Coast Reserve LTER

Annual Data Managers' Meeting Scheduled

The Annual LTER Data Managers' meeting is tentatively scheduled to start Thursday evening, August 1 and conclude on Saturday afternoon, August 3. It will immediately precede the ESA events that begin on Sunday, August 4 in San Antonio, Texas.

Discovering NASA Data

The National Aeronautics and Space Administration (NASA) has a diverse collection of data systems which provide information on everything from satellite images to global vegetation maps and models. The individual systems are scattered around the country at different NASA headquarters. Fortunately, there is a unified system, known as the Master Directory, that can provide information about what database systems are available and link you to them.

The Master Directory can be accessed via the Internet, regular phone lines or the commercial Telenet network (see below for details). The Master Directory program provides options that give you information on over 15 separate NASA databases. This information includes: summaries of data sets, "campaigns," sensor platforms (both aircraft and satellite), individual sensors and data set archives.

The most powerful (and useful) feature of the Master Directory is its ability to directly link to almost all of the listed databases. You can jump directly from a description of a database to the database itself (regardless of where in the country the database actually resides). Doing this is as simple as selecting the LINK option from a menu.

Individual databases are as varied as the scientific community NASA is trying to serve. Of particular interest to terrestrial ecologists are the Earth Sciences Data Directory (an interactive directory of over 1,000 data bases), the Earth Resources Observation Systems Data Center, the Pilot Land Data System (a compendium of NASA controlled imagery), the European Space Agency/Earthnet Program Office Catalog and the Synthetic Aperature Radar Data Catalog, or to put them in NASA terms, the USGS ESDD, the EROS-EDC, the PLDS, the ESA-LEDA and the SDCS.

The use of acronyms and sheer mass of data sources pose the most serious obstacles to the successful use of the data systems. Also once you get past the Master Directory into individual data systems, the

quality of the user interface varies widely. Many are based around front-ends to SQL (Structured Query Language) databases. Several have graphical interfaces, when used on terminals that emulate graphics terminals such as the Tektronix 4014, that aid in selecting spatially specific data. All of them incorporate esoteric elements that will take some practice to master.

Virtually all the databases in the Master Directory are set up to provide meta-data (data about data) but not the data themselves. This is understandable considering the huge size of many of the individual datasets and images. Most of the databases provide a mechanism for ordering data.

Using the NASA Master Directory

To access the Master Directory, you will need a computer terminal (preferably one that emulates a VT-100 or ANSI terminal). Below is listed the relevant connection information for logging on from several different networks. No prior registration is required to use the Master Directory.

Internet: Telnet to NSSDCA.GSFC.NASA.GOV (or 128.183.10.4)
Use the username NSSDC

Direct Dial: Set modem to 8-bits, no parity, 1-stop bit, 300-2400 baud. Dial 301-286-9000. Enter number MD and username NSSDC.

SPRINTnet (Telenet): At the @ prompt: 32107035/GSFC, use password: 0356156 and username NSSDC.

More help may be obtained via electronic mail to: MDSUPPORT@NSSDC.GSFC.NASA.GOV

GIS Corner

Report Issued-- An LTER report entitled "Technology development in the LTER Network: Current status of GIS, remote sensing, Internet connectivity, archival storage and global positioning systems," by David Foster and Emery Boose, was completed in March and is planned for publication by the LTER Network Office as part of its new numbered series. The report incorporates the results of last summer's survey of GIS and remote sensing capability across the Network. The assistance of James Brunt, Bill Michener, Rudolf Nottrott, and John Vande Castle in assembling this report is gratefully acknowledged.
--Emery Boose, Harvard Forest LTER

ARC/INFO and USGS DEM's: Happy day. The USGS tape with 15 7.5 minute DEM's had arrived. We didn't have a 9-track tape drive so arrangements were made with the Computer Engineering department to read the files off the tape for us. And in a couple of days the data was on our system. The Tin Manual was read from cover to cover and the pertinent sections, the Quick tour, Processing TIN data, and the Appendices D and E, were studied. With a crackle of knuckles we were ready to begin. The command
demlattice dem01 dem01.lat was carefully typed in and...
Loading header information...

Premature end of file found while reading dem header (DEMHDR).
FATAL ERROR

Unable to convert dem to lattice (DEMLAT).
Bailing out of DEMLAT
Arc:

What!? The command demlattice dem01 dem01.lat was carefully.....   No
go. What the sam hill is going on? The manual is studied again, the
files are compared, and permissions are checked. The file is compared
with the Appendix D listing of the logical record format of USGS 7.5
minute DEM's. The file is exactly right, however there are no hard
returns in the entire file. Have you ever tried to edit a 1190748
byte file with no hard returns in Emacs?

With a little searching the example DEM quick.dem is found in
/research/esri/samples/tin. This file is compared with our file. The
quick.dem sample has hard returns. Hard returns are added to the end
of each logical record using a c program.

The command demlattice dem01 dem01.lat was carefully.....
Loading header information...
DEM extent [315557.094, 3777408.750, 327341.375, 3791490.500]...
DEM surface range [1401.000, 2192.000]...
Sample distance in x and y [30.000, 30.000]...
Loading 1:24000 DEM data into lattice...
(Hey, All Right!!!)
End of file found while reading dem (DEMCON).
FATAL ERROR
Unable to convert dem to lattice (DEMLAT).
Bailing out of DEMLAT
Arc:

What!!?? *+(#$%

Various machinations are tried, the hard returns are checked, the c
program is checked. Everyone goes home and I think of logical records
A and B, and of profiles, true north and UTM north during dinner.

ESRI is called. "No one is available at this time, may we call you
back?" Yes. The next day they call. I try to explain about logical
records and hard returns. "May I refer you a technical person that
knows TIN". Yes. The next day they call. What? My fault? Data
corrupted or in the wrong format? "Have you tried the TIN example
quick.dem". Yes. "What is the exact setup of your hardware and soft-
ware". DECstation 3100, ... "Oh. You don't have a 9-track local to
your machine". No. "How did you read the DEM off of the tape".
Computer Engineering, etc .... "Did you use demread?" What?

"DEMREAD, located in /esri/arcexe50/utool". Utool directory? "Yes".
Well I read a little discussion in Chapter 10 in the AML manual about
the atool directory and a word or two about utool there, and in
Chapter 13, Advanced functions in ARC/INFO, and I read and reread the
sections on DEM's in the TIN, but nothing ANYWHERE about demread.   Of
course I could have looked up DEM's in the index, but as you know
there is no index. Maybe in the next release.

Ok,

```
> cd /arcexec50/utool
> ls
config*  config.list*  demread*  remove*
```

Well, sure enough!
```
> cat demread
case $# in
0) echo "USAGE: demread <filename>" ;;
1) dd if=/dev/nrmt0 of=$1 ibs=1024 cbs=1024 conv=unblock
esac
```

Hey, How about that, cbs=1024 conv=unblock. It's putting hard returns at the end of every 1024 byte and unblock is converting fix length records to variable length records. Hey, why didn't I think of that!

So, the original files that came off tape with the command:
dd if=/dev/nrmt0 of=dem01 ibs=1024

were run through the command:
dd if=dem01 of=dem001 ibs=1024 cbs=1024 conv=unblock

and ....
The command demlattice dem01 dem01.lat was carefully.....
Loading header information...
DEM extent [304606.500, 3805353.750, 316375.781, 3819451.500]...
DEM surface range [1587.000, 2767.000]...
Sample distance in x and y [30.000, 30.000]...
Loading 1:24000 DEM data into lattice...
Writing out lattice...
Output lattice is 392 points in x, and 470 points in y.
Arc:

Nothing to it!

Now, all we need to is put then all together and shazam instant surface.   Sorry, wake up, stop dreaming. Like the man said, nothing is ever simple. And like the that other guy said, Just the facts ma'am, just the facts.

In Arc/Info there is a size restriction on both a lattice and a tin. A lattice cannot be larger than 1024 by 1024 and a tin cannot have more than 50,000 points in it. The Sevilleta LTER needs 15 7.5 minute USGS quadrangles to cover its extent. A single quad sampled at 35% using VIP results in 60 to 65 thousand points being selected.   In addition, specifying a Z tolerance of 10 meters can take more than 45

minutes and then exceed 50,000 points anyway. For creating maps of slope and aspect the accuracy of the original DEM's are usually required, so using a TIN is ruled out from the start. The solution thought of, and approved of by ESRI, was to create slope and aspect maps from the lattice file using LATTICEPOLY from each individual lattice (corresponding to each DEM). And then MAPJOIN the resulting coverages together. This entails some inaccuracy of the polygons at the edges of each area but given the extent of the whole area and the proportion to the area affected, the degree of error introduced is thought to be acceptable. It would be possible to eliminate this error by creating polygon coverages that produce overlap between individual lattice files, join adjacent lattices, LATTICECLIP these

lattices with the clip polygon, create a slope or aspect map, clip these maps with the original boundaries of the quads and finally MAPJOIN the result. However, in the estimation of the person I spoke to at ESRI, this effort is for a minimal gain, especially given that we wanted to have slope and aspect in five degree increments, which would result in fairly small polygons so the actual area affected would be smaller than the normal 45 degree chunks that ESRI uses as an example.

However, for producing a surface model of your LTER site, if it is more than four quads in extent, there is no work-around. There is nothing to do but run VIP on each lattice file so that the resulting lattice that is created by LATTICEMERGE does not exceed 50,000 points. And if a TIN is required, set the Z tolerance at a very low number and increase this until ARC can successfully create a TIN with the specified tolerance. This requires that VIP be run with the histogram option for each lattice so that the exact number of points selected is known. In our case, a polygon clip coverage was made to reduce that area to that minimum which would encompass the extent of the LTER. Of course, the real solution is the software company's panacea, ... the next release. Which is now, according to the person I spoke with, version 6.0, which will remove both restrictions, since both lattice and TIN data structures will be disk based. --Tom Garrison, Sevilleta LTER

Ask Dr. Network

Q: I was attending a meeting at Oregon State. They were able to arrange for us to phone into one of their computers and use telnet to get to our own computers. It made it possible for us to complete a NSF proposal on time. Can other sites could offer visiting LTER scientists an account name which they could use temporarily to communicate with their home base?
A: Since close to 90 percent of all main and associated locations have an Internet connection (or are working on getting one), most of those

sites would be technically capable of giving you a temporary account for communication with your home base. Based on experience, however, university administrations are not always as flexible as OSU in providing temporary short-term accounts, so you'd probably end up 'borrowing' the account of somebody else at the site.

As a more general solution to this problem, the network office staff is presently working with SPRINTnet (formerly TELENET) to get a SPRINTnet connection to LTERnet installed at the Network Office (as recommended in the connectivity report). When this is in place, any LTER member can dial a local SPRINTnet phone number to access the LTERnet host computer, which is on the Internet. Then, when you are away from home visiting a place that cannot do what OSU did for you (example: the hotel at Snowbird during the ESA meeting), you can always dial LTERnet via a local SPRINTnet number and from there Telnet to anywhere on the Internet. Since your account on LTERnet, once established, will be permanent, it will always look and behave the way you set it up.

Q: A connection to Sprintnet sounds great. Let me know when it works and how I get to be a member of Sprintnet.

A: The LTER Network Office will keep you posted. You don't need to become a member of SPRINTnet, we'll take care of that through the Network Office. All you will need is access codes, and the LTERnet password for your account.

Scientific Data Management Reports

Reports on a NSF workshop on scientific database management in March of last year are now available via anonymous FTP from LTERNET. The reports make numerous recommendations, most of them of direct interest to LTER data managers. Several of the recommendations have already been implemented within the LTER network. The full reports provide much greater detail on the recommendations, but a summary is listed below.

Professional societies should:

1) Promote standardization of data interchange descriptions that use self-describing data formats.

2) Standardize description of data within each scientific discipline.

3) Require appropriate citation of data and the deposit of relevant data into the appropriate archive before permitting publication in the societies' journals.

4) Promote and fund workshops to investigate and recommend policy relative to the retention of data.

5) Promote and fund workshops directed towards resolving the sociological problems hindering the development of sound data management policy.

NSF should lead the scientific community to:

6) Perform research in methods for describing metadata and promote standardization in the management of metadata.

7) NSF should include the matter of cataloging and publication of databases in planning of NSFNET and other national networks. Research proposals for projects leading to databases should be required to include plans for publishing, cataloging, and either maintaining or transferring results to national archives.

8) Perform further basic research in storage device technology and in representation techniques for better utilization of the available capacity.

9) Perform basic research in information science and information retrieval to improve the ability of interested researchers to locate relevant scientific data.

10) Support the creation of one or more data analysis environments for science data that includes database and data archival processing. In these environments, special attention should be given to integration of the user interface, the analysis component, and the database

management system.

11) Perform basic research in emerging database technologies such as object-oriented database systems, extensible database systems, and logic database systems. Further, explore alternatives to the relational model of data such as models directly supporting lists, sequences and graphs.

12) Perform research directed at solving the problems of heterogeneous data management environments.

13) The NSF should coordinate an multiagency task force charged with promoting the creation of a harmonious environment which enhances the ability of scientists to freely and easily exchange data.

The reports are available in both Postscript and text forms on LTERNET

and I highly recommend them to LTER data managers. --John Porter, Virginia Coast Reserve LTER

Announcements

Getting the Word Out: At the Andrews we have found that offering a quarterly QSG (Quantitative Science Group) workshop series is an efficient means of technology transfer. This series is highly subscribed and very popular with our students and faculty. We offer these sessions on a first-come, first-served basis. They are free to members of our Department and colleagues associated with the AND LTER. Other interested participants are charged a nominal fee and are admitted as space permits. Here is a listing of this quarter's offerings: Introduction to the Forest Science Network, Advanced Network Features, Introduction to the Sun Workstation, Introduction to Geographic Information Systems (GIS), Introduction to ArcInfo, Advanced ArcInfo, Introduction to Unix, Bravo Slide Maker, Remote Login and Internet Communications. --Susan Stafford, Andrew's LTER

Monitoring Workshop: Oregon State University will be hosting a workshop entitled: "Improving Natural Resource Management Through Monitoring." The workshop is hosted by Susan Stafford of the Andrew's Forest LTER. Workshop speakers are drawn from a wide array of private an public institutions, both large and small. It will feature a wide spectrum of monitoring programs, with a heavy emphasis on the data management resources needed to implement successful monitoring. The first day will feature an overview of monitoring programs (with examples). The second day will cover elements of design for monitoring programs (during this session Susan will be speaking on the "Nuts and Bolts of Managing Research Information" and John Vande Castle of the LTER Network Office will be speaking on "Networking Issues and Connectivity"). The final day will focus on the topics of future technology and program evaluation.

Science and Technology Information System: NSF has a new public-access information system. It can be accessed over the Internet (telnet to STIS.NSF.GOV or 128.150.195.40) at no charge. To log on, you enter 'public' at the login prompt. It will then ask you for a terminal type. Supported terminals are VT-100 and Sun. I recommend using the VT100nkp (short for, VT-100, no keypad) if you are using the NCSA

TELNET program from a PC or Mac. The regular VT-100 emulation ignored my arrow keys when using NCSA TELNET and arrow keys are definitely needed to run STIS. Full help on STIS can be obtained via anonymous FTP to STIS.NSF.GOV where you can download STIShelp.man (it is ASCII, a WordPerfect version is also available).

It is very easy to use STIS (provided that your terminal emulation is correct). It supports a fairly advanced user interface, with pop-down menus and scrolling windows. Performing searches on the 400 or so documents indexed is relatively easy, with options to display search results on the screen or download the text for local printing. The program responds rapidly (at least when the net is clear!) and provides an excellent way to obtain information about a wide range of NSF programs. If you lack easy access to the Internet, you can also reach STIS by phone (use 300-2400 baud, even parity, full duplex and 7 data bits) at (202) 357-0359 or 357-0360. A copy of the STIS reference card came out with the March-April issue of the LTER Bulletin. -- John Porter, Virginia Coast Reserve LTER

Finagle's Rules

Ever since the first scientific experiment, man has been plagued by the increasing antagonism of nature. It seems only right that nature should be logical and neat, but experience has shown this is not the case. A further series of rules has been formulated, designed to help man accept the pigheadedness of nature.
Rule 1: To study a subject best, understand it thoroughly before you start.
Rule 2: Always keep a record of data -- it indicates you've been working.
Rule 3: Always draw your curves, then plot the reading.
Rule 4: In case of doubt, make it sound convincing.
Rule 5: Experiments should be reproducible -- they should all fail in the same way.
Rule 6: Do not believe in miracles -- rely on them.
TECHNICAL FORUM: Designs for Distributed Databases

Editor's note: Integrating the individual LTER site databases into a network context is technically challenging. This section includes articles describing different approaches to development of special purpose, network-wide databases.

Network Access to Data Using Electronic Mail

Last July, at their annual workshop, the data managers discussed the possibility of developing an inter-site climate database, to provide a means by which researchers at any LTER site can obtain climate data from other sites. They agreed that the time was right to develop a prototype, one that would start with a few sites and later be extended to others would also be extended to include other types of ecological data.

The Kellogg Biological Station (KBS), along with Bonanza Creek (BNZ), Central Plains (CPR), Konza Prairie (KNZ), and Northern Temperate Lakes (NTL) recently proposed to develop such a prototype, to be funded as an Inter-site Technology Supplement. The network office also will take part.

There will be two methods of access to the proposed database. An "email-gatherer-server" system will be developed at KBS, and software for "interprocess communication" will be developed at CPR. These two methods, though they overlap somewhat, have different, complementary purposes. A gateway between the two will be set up at KNZ.

I describe here a few details of how the email-gatherer-server is intended to work.

Each site will continue to manage its own climate data locally, but you will be able to send an email message with your request to one central address, and receive the requested data via the network. Under the prototype, only data from BNZ, CPR, KBS, KNZ, and NTL will be available, but any LTER researcher with access to email will have access to the data. Later on, it is hoped that other sites will be able to make their data available, too.

The name "email-gatherer-server" refers to the fact that a central "gatherer-server" computer will gather data for you from several sites, and serve them to you via email. The gatherer-server will be able to reconcile different formats and representations of data, combining them into a single usable format to be sent to you via the network.

It will also be able to summarize specialized climatological variables much more conveniently than any statistical, database, or spreadsheet software can now do. These functions are intended to help overcome some of the technical hurdles, each of which may be relatively minor in itself, that taken together constitute a serious obstacle data sharing.

The exact details of the "language" by which you will be able to request data will be worked out as part of the proposed work, but our initial thoughts are along these lines: To request data, you will send an email message to data@lternet.washington.edu (or some such Lternet address). You will be able to get more information about the database by sending a message consisting of one word, "Help." To find out what climate databases (and variables) are available, you will be able to send a message like "show sites" or "show variables." To get specific information about, for example, what precipitation data are available, you will be able to send a message like "show precipitation" or "show rainfall." To get actual data, you will be able to send a "send me" message, specifying the variable or type of variable, dates, units of measurement, summary interval (e.g. daily, monthly, or 30-year average), and data format you want.

Since email cannot give you instant responses to your requests, we will design the system to reduce the number of trial-and-error

iterations needed. If you are not specific in your request, you will be sent a sample set of data, along with information to help you make a more specific request. For example, if you ask for monthly precipitation data, but don't specify what months or what sites you are interested in, you will be sent a few months worth of several sites' precipitation data.

With these functions, the proposed system promises to be a major step

forward in LTER data management in support of collaborative, inter-site research. --John Gorentz, Kellogg Biological Station LTER

Facilitating Access to Data Using Interprocess Communication

Fulfilling a request for data requires a significant expenditure of personnel time at most sites for reasons such as security, ease of access to the data by local personnel, the computer systems involved with data storage, and the nature of the database software or system being employed. One of the goals identified by the LTER Data Managers at their last meeting was to develop mechanisms to promote the exchange of data between sites. John Gorentz proposed developing a batch-oriented system using electronic mail as the basis for data exchange, because E-mail is almost universally available. I proposed that we consider developing the software tools needed to allow interactive access to the LTER databases using a layer of network protocols below the E-mail level. John recently took the initiative to organize a multi-site proposal to NSF to pursue developing a prototype distributed database system. The proposal was submitted with the title of "Inter-site Climate Database for Agricultural and Ecological Research (LTER Technology Supplement)", under the leadership of G.P. Robertson, M.J. Klug, and E.A. Paul (Michigan State University/Kellogg Biological Station.). For this issue of DataBits I have extracted from my part of the proposal some of the material that describes what we would like to accomplish.

Interprocess communication (IPC) under the UNIX operating system can be used to allow two or more processes to execute concurrently and exchange data. The processes may execute on different computers, or on a single multitasking computer. Information exchange between the processes is handled using network IPC functions. The data exchanged using these techniques are not limited to ASCII files, nor are they limited by length, as are data files passed using electronic mail. Binary data files or files of GIS images can be exchanged as easily as sequential access ASCII file.

Interprocess communication under UNIX can be handled in a variety of ways. Traditional IPC mechanisms under UNIX include "pipes", a one-way

mechanism for passing data, and "socketpairs", which corresponds to a pair of pipes and allows two-way communication. These mechanisms have the limitation that they can only be used with "children" of a single "parent" process. This limitation means that pipes cannot be used to link two or more independent processes. In addition, pipes cannot be used between processes on separate computers.

Recent versions of BSD UNIX support IPC techniques that permit communication to be established between independent processes running on either one computer or across two or more computers. These mechanisms rely on network "sockets" for communication, and can pass data either as "datagrams" or as "streams". Datagrams are discrete messages that are transmitted by one process to another. Neither the receipt of the datagram nor the order of delivery is guaranteed with datagrams, so the communicating processes are required to establish a protocol for passing messages. Data streams do insure delivery of data in the order in which it was sent, but at a somewhat higher cost in terms of overhead associated with the data transmission. Data streams

can be opened either in the UNIX domain, to exchange data on a single computer, or in the Internet domain, to exchange data between computers.

We have begun to explore the use of IPC facilities under UNIX to improve access to our database. Our interests are in making it easy for people to access and view the data, and to make it easy to write models that can access the data in a relatively transparent manner. We recently proposed to extend our development effort to help facilitate data exchange between LTER sites.

We are using a client-server paradigm in our design of software. The server is a program that services the needs of one or more client programs. The server would accept requests for data from the client process and attempt to fulfill them. Across the LTER network the servers would be customized to extract data from the local database of the cooperating sites. The customizing of the servers would be required because of the diverse nature of the databases within the LTER network. The client program acts as an interface between a scientist and the computer systems. The client would provide the ability to view data in tabular or graphical formats on Sun workstations, and to transfer data from one site to another.We also proposed to extend our work by providing similar capabilities for accessing data from IBM compatible PCs using the PC-NFS functions.

Our primary goal is to provide a library of functions that can be used with either Fortran or C to easily build server and client programs that can communicate using the BSD UNIX stream sockets for interprocess communication. The interface to Fortran programs will be provided because many of our models are written in Fortran, and

because many sites do not have proficient C programmers. The library will be the basis for tools designed for interactive and non-interactive transfer of data. The interactive transfer will include functions to display data as tables or graphically under SunView and on PCs, and to capture the data in local files. Graphics routines that run under SunView also will be included in the library.

It may sometimes be more convenient to transfer sets of data in an unattended mode. We also proposed to develop functions and procedures for scheduling file transfers using RCP, FTP, or other file transfer protocols as non-interactive processes. These will be designed for the transfer of data files in a batch oriented mode.

We plan to make available an example system that uses our locally developed file description system for accessing sequential access files. We plan to use this mechanism to create a single server that can access any of our publicly available data sets. Our file description system can be easily integrated with a database server for our local files because it provides the information needed by the server to read the data, such as field types and locations on the data records. We plan to provide access to the data descriptions as well as the data, allowing potential clients to "browse" our database to learn what is available.

I have written a server that can access our one set of our climate files, and a client that can request the data from the server. I am in the process of getting it tested, and then will make it available as

an example of what can be done. I am planning on adding some graphics capabilities to it as well, to give you some examples that may be of use for writing your own graphics routines under SunView. If you are interested in getting this example please contact me. I will make it available through anonymous FTP. --Tom Kirchner, Central Plains Experimental Range LTER

Distributed Herbarium Database

A milestone in the networked herbarium database project at KBS was reached recently. Although this is not an LTER project, the herbarium database is one of our core LTER datasets, and some ideas for the recently proposed inter-site climate database grew out of this project.

In April 1990, we developed a working prototype, by which a person using software on the KBS VAX could obtain data not only on the 6000+ specimens in our own herbarium, but also on specimens from the 18000+ specimen Borneo Collection in MSU's Beal-Darlington Herbarium. While the Ingres DBMS is used to manage both collections, the databases are

on two different computers at two separate locations. The KBS Collection is on our own VAX/VMS system, and the Borneo Collection is on a Sun SparcStation on the main MSU campus. One can query both databases using the same set of menus. Behind the scenes, queries for Borneo data are sent to the Sun SparcStation by email, and the data matching the request are received by email.

The software developed last year used several VAX/VMS specific features, and queries could be issued only from the KBS end. Recently the software was converted to a portable version that runs on the Sun Sparcstation as well as on our VAX/VMS system. Either database can now be queried from either location.

Our next step, already underway, is to develop a mechanism by which anyone with access to Internet email can query the databases without the need for any local database software. The collections are being indexed on every word in the databases, much like many computerized library card catalogs are now indexed on every word on the catalog entries. This means you will be able to do a keyword search on any term. For example, to obtain information on poison ivy specimens collected in Borneo by J. Doe, you could send an email message addressed to the Beal-Darlington herbarium, containing the line, "k=Toxicodendron and Borneo and Doe." The data on the matching specimens would then be sent to you by email.

The software development is being done by Tim Seeley, a student at Western Michigan University, and is directed by Stephan Ozminski. It is funded by an internal MSU research initiation grant to John Beaman and Jim Beach of the Department of Botany and Plant Pathology. --John Gorentz, Kellogg Biological Station LTER

Andrew's LTER Dataset Use

Here is a table of the most recently requested datasets in the Forest Science Data Bank over the last year or so. An additional 11 data sets were requested by a single individual. We have been pleased with the broad geographic and institutional interest as well as the diversity of uses these datasets have been intended for.

Recent Requests of FSDB Data and their Uses

Code	Title	Uses	User	Institution
CF02	HJA Stream Water Chemistry	Nutrient Chemistry Intersite Variability Court-ordered Evaluation	Martin Magnusson Beschta	HBR * NTL * OSU

GS01   Stream   Analysis of Changes in   Grant   AND
GS02   Channel   Channel cross-section
GS09   Cross-   Geometry
GS11   Sections
HF04   HJA   Rain on Snow Model   Harr   USFS/UW
       Streamflow   Forest Biogeochem. Model   Running   U. of
Montana
                     Streamflow Model   Hicks   New Zealand
                     General Analysis   Stednick   CSU
                     State Data Collection   Robison   OR Water
Res.
                     Intersite Variability   Magnusson NTL *
                     Sediment Yield Model   Grant   AND
                     Rainfall-Runoff Model   Grant   AND *

MS01   Primary   Forest Biogeochem. Model   Running   U. of
Montana
       Meteorol.   Climate Model   Russel   ARC
       Station,   Water Balance Model   McCoy   U. of Mass.
*
       HJA   Climate Model   Webb   OSU
                    Rainfall-Runoff Model   Grant   AND *
MS02   Climatic   Forest Biogeochem. Model   Running   U. of
Montana
       Station at   Rain on Snow Model   Harr   USFS/UW
       WS2.   Climate Model   Russel   ARC
                    Water Balance Model   McCoy   U. of Mass.
*
                    Climate Model   Webb   OSU
                    Rainfall-Runoff Model   Grant   AND *
                    Sediment Yield Model   Grant   AND
TP09   Destructive Forest Succession Model   Urban   VCR *
TP14   Biomass   Garman   COPE/OSU *
TP15   Data   Sollins   AND *
TP18   Spies   AND
                                                Harmon   AND *
TV005 Charact.   Remote Sensing Model   Cohen   AND
       Old-growth
       Forests
TV010 Forest   Forest Succession Model   Urban   VCR *

       Reference   Garman   COPE/OSU *
       Stands   Sollins   AND
                                                Spies   AND
                                                Harmon   AND *
                    Remote Sensing Model   Cohen   AND

* = For use in inter-LTER-site analysis

LTERNET Mail Forwarding Groups

All addresses are suffixed by @LTERNET.WASHINGTON.EDU

AERC - Association of Ecological Research Centers bBedford, jGosz, jHobbie, gLikens, oLoucks, jReynolds, fWagner

ANT - Atarctic Palmer Station LTER Site; see PAL wFraser, rHanson, eHofmann, pPenhale, lQuetin, rRoss, rSmith, wTrivelpiece,

BRIE - Biogeochemical Reactions in Estuaries LMER site tHollibaugh, wKimmerer, fMackenzie, hPaerl, fSansone, StephenSmith

CRE - Columbia River Estuary LMER site jAnderson, jBaross, dJay, dReed, cSimenstad, lSmall, bWissmar (Usernames NOT on the list because we don't have valid e-mail forwarding addresses for them: fPrahl, dMcIntire)

climate - LTER Climate Committee pAdams, eBoose, jCrum, cDahm, tFederer, dGreenland, bHayden, dHill, jHobbie, wJarrell, tKittel, bKjerfve, bLawrence, jMagnuson, aMckee, bMichener, tSeastedt, rSmith, lSwift, lViereck (Usernames NOT on the list because we don't have valid e-mail forwarding addresses for them: jTester, wWendland)

cohort1 - representatives of the Cohort1 LTER sites: nCaine, dKaufman, wLauenroth, jMagnuson, jMeyer, fSwanson, jVernberg

connectivity - LTER Connectivity committee and its advisors: cBledsoe, jBrunt, rNottrott, jPorter, rRobbins, dVanBelleghem

datamnt (or datamng, or dman) - LTER site data managers and people interested in data management issues: pAdams, bBenson, cBledsoe, eBoose, cBowser, jBriggs, jBrunt, gCalabria, eChapal, aElhaddi, jGorentz, jHalfpenny, dHenshaw, rIngersoll, tKirchner, mKlingensmith, mKlopsch, jLaundre, dLightfoot, eMelendez, bMichener, eMuldavin, rNottrott, sOzminski, jPorter, rRobbins, tSabin, gSpycher, sStafford, cVeen

datatask - LTER data managers' task force: jBrunt, bBenson, bMichener, rNottrott, jPorter, sStafford

decomp - users interested in the subject of "decomposition", and related issues (intersite decomposition experiments, etc.): jAber, cBledsoe, lBlum, rBoone, lBoring, iBurke, rDueser, bEdwards, tFahey, tGower, mHarmon, sHart, lLagera, jMacMahon, jMelillo, jMeyer, jMorris, kNadelhoffer, bParton, jPastor, ePaul, jReynolds, aRicca, tSeastedt, gShaver, tSiccama, pSollins, kVanCleve, mWalker, dWedin,

cWhite, wWhitford, (Usernames NOT on the decomp list because we don't have valid e-mail forwarding addresses for them: hGholz, jLodge)

gis (or giswork, or gisworkshop) - users interested in GIS-related issues: pAdams, lAshkenas, sBledsoe, eBoose, jBriggs, iBurke, jCallahan, bCunningham, gCunningham, sDegloria, rDueser, dFoster, jFranklin1, sGage, jGorentz, sGregory, dGrigal, dHall, hHammond, rHoffer, rIngersoll, bKjerfve, tKratz, lKrievs, lLagera, bLauenroth, jLaundre, lLestak, gLienkaemper, mMackenzie, jMeyer, bMichener, rNottrott, rParmenter, jPorter, kSaari, tSchwartzman, tSeastedt,

sSmith, dStow, dTilman, dTomlin, jVandeCastle, lViereck, rWaide, rWynne, jYarie (Usernames NOT on the gis list because we don't have valid e-mail forwarding addresses for them: kshaw, dhall, dwalker)

equip - for people interested in the acquisition of software, image data and hardware, as related to GIS and remote sensing: jAber, pAdams, bBenson, cBledsoe, eBoose, cBowser, jBriggs, jBrunt, iBurke, nCaine, gCalabria, jCallahan, wCohen, dCrossley, jCrum, bCunningham, gCunningham, sDegloria, cDriscoll, rDueser, aElhaddi, jFranklin, jFranklin1, sGage, jGorentz, jGosz, dGrigal, bHaines, jHalfpenny, hHammond, dHenshaw, jHobbie, dHall rHoffer, rIngersoll, dKaufman, tKirchner, bKjerfve, mKlopsch, tKratz, lLagera, bLauenroth, jLaundre, wLawrence, lLestak, dLightfoot, mMackenzie, jMagnuson, rMartin, eMelendez, jMeyer, bMichener, bMusick, rNottrott, sOzminski, rParmenter, dPeterson, jPorter, rRobbins, gRobertson, kSaari, tSabin, bSchlesinger, tSchwartzman, tSeastedt, sSmith, gSpycher, sStafford, dStow, fSwanson, dTilman, dTomlin, kVanCleve, jVandeCastle, cVeen, jVernberg, lViereck, rWaide, cWessman, wWhitford, sWalker, rWynne, jYarie (Usernames NOT on the list because we don't have valid e-mail forwarding addresses for them: dwalker)

exec - LTER Executive Committee: cBledsoe, jFranklin, jMagnuson, jMeyer, bSchlesinger, kVanCleve, jVandeCastle,

global - global change workshop participants: fSwanson, eBlood, dFoster, jMagnuson, tSeastedt, jGosz, dCorrell bLauenroth, lViereck, jHobbie, jMelillo, cDriscoll, nCaine, gRobertson, bHayden, jReynolds, pRisser, rWoodmansee, rWaide, aLugo, dGreenland, wSwank, dFoster, sTarapchak, rSharitz

hydro - Hydrology Working Group mAnderson, dArmstrong, cAsbury, cBledsoe, eBlood, sBolin, bBowden, cBowser, iBurke, nCaine, jCornelius, jCrum, cDahm, rDolan, cDriscoll, tDyrness, tFederer, dFoster, jFanklin, lGardner, eGorham, gGrant, dGreenland, sGregory, cHall, bHayden, gHornberger, jKoelliker, gLikens, mLitaor, wMartin, aMcKee, jMeyer, mMolles, rNeilson, wNuttle, rPierce, pSollins, wSwank, fSwanson, cTate, bWallace, jWebster, tWilliams, rWoodmansee (Usernames NOT on the list because we don't have valid e-mail

forwarding addresses for them: kBrooks, mDavis, dDecoursey, dDoehring, aGiblin, gKipphut, mMcElroy, fScatena, tWard)

LMER - Land Margin Ecosystems Research Group: BRIE, CRE, PROTEUS, WAQ, pTaylor, lDuguay

LMERPI - Principal Investigators of the Land Margin Ecosystems Research Group: cSimenstad, dJay, tHollibaugh, stephenSmith, rCostanza, mKemp, jKremer, iValiela, pTaylor, lDuguay

modelers - Ecosystems Modelling Group pBacon, dCoffin, bLauenroth, rSantore, pSollins,

net - everybody at the LTER Network Office: jfranklin, cbledsoe, rnott, smartin, jvandecastle

PAL - Antarctic Palmer Station LTER Site: wFraser, rHanson, eHofmann, pPenhale, lQuetin, rRoss, rSmith, wTrivelpiece,

PALEO - Paleo-ecology group dArmstrong, sElias, dFoster, tFrost, dGoodreau, bHayden, tKratz, kWilliams, dYamaguchi (Usernames NOT on the list because we don't have valid e-mail forwarding addresses for them: wEisner, jEllis, kEveritt, wPatterson, dShell)

pi - LTER Principal Investigators (extended list for site information): jAber, cBledsoe, eBlood, lBlum, eBoose, cBowser, jBriggs, jBrown, jBrunt, iBurke, nCaine, dCrossley, gCunningham, cDriscoll, tFahey, nFetcher, dFoster, jForwood, jFranklin, eGorham, jGosz, dGrigal, jHalfpenny, bHayden, jHobbie, rIngersoll, dKaufman, mKlug, tKratz, bLauenroth, aLugo, jMagnuson, sMartin, aMcKee, jMeyer, wMichener, bMoller, rNottrott, ePaul, jReynolds, gRobertson, bSchlesinger, tSeastedt, gShaver, hShugart, pSollins, wSwank, fSwanson, dTilman, jTorrey, kVanCleve, jVandeCastle, cVeen, jVernberg, lViereck, jWelch, rWaide, aWhitener, wWhitford, PAL

PROTEUS - Processes of Recycling, Organic Transformation and Exchange between Uplands and the Sea within Chesapeake Bay (Chesapeake LMER site): rCostanza, mKemp

remote - users interested in the subject of remote sensing and related issues: jAber, cBledsoe, eBoose, jBriggs, wCohen, dCrossley, jCrum, gCunningham, sDegloria, jFranklin1, dHall, lLagera, rLathrop, wLawrence, tLillesand, mMackenzie, rMartin, bMusick, rNottrott, dPeterson, dStow, dTomlin, lViereck, cWessman, rWynne, jVandeCastle
roots - Participants in the root analysis workshop, June 1990, MSU Lansing, Michigan blauenroth, jbattles, lblum, rfogel, sgower, hpersson, kpregitzer, rruess, asmucker, jflore, cbledsoe (NOT on the

root list because we don't have valid e-mail forwarding addresses for
them: pHook

SQL - Structured Query Language (SQL) implementation group bBenson,
jBriggs, jGorentz, rNottrott

WAQ - Waquoit Bay LMER site dAubrey, cDAvanzo, kForeman, jKremer,
kLajtha, pPeckol, cSham, iValiela (Usernames NOT on the list because
we don't have valid e-mail forwarding addresses for them: gGeiser)

DATABITS Publication Information