LTER DataBits
Information Management Newsletter of The Long Term Ecological Research Network

01001100 01010100 01000101 01010010
01001100 01010100 01000101 01010010

◆ *Feature Articles*

DataBits Newsletter New Design
NT Networks
NACSE-LTER Funding
Network Office Exchanges

◆ *News Bits*

Database Design Tools
Web Survey and Catalogue
Electronic Authoring
Network Office Updates
Software Bits

◆ *Calendar*

Calendar Events
LTER Education Workshop 22-24Oct98
Data Manager DataTask Meeting 18-23Feb99
LTER Technology Committee Meeting 26-27Mar99
Data Manager Meeting 04-06Aug99

Feature Articles

LTER Newsletter DataBits New Design

- Karen S. Baker, Palmer LTER and Marshall White, LTER Network Office

DataBits, the LTER Data Manager newsletter, has a new design. The newsletter was developed in 1990 as a mechanism for the LTER sites to share information. Publication continued through 1994 under the editorship of John Porter (VCR). A visit by Karen Baker (PAL) to the network office catalyzed the Network Office plan to recreate DataBits. A prototype was developed for presentation to DataTask and was built upon the concepts of modular design, online publication, co-editorship as well as both inreach and outreach opportunities. The redesign for Spring99 DataBits was developed in conjunction with Marshall White (NET), a part-time student graphic designer, who works with Patty Sprott (NET) to produce the online version of the LTER Network News. The initial guiding principles listed below will evolve as the concept of rotating editorship develops.

- Keep layout simple and modular
- Make look coherent with LTER Network News and LTER Web
- Design for online presenation, but store as single file for simple hardcopy output
- Promote cross-site and cross-network data manager communications
- Co-ordinate editor responsibility by having rotating co-editors
- Facilitate remote editing (ie Netscape Composer is free)
- Develop in working documents location to facilitate co-editing prior to posting
- Provide Table of Contents for quick browsing
- Provide summary box to serve as email announcement with active hyperlink
- Sections to include Summary Box, Feature Articles, News Bits, and Calendar
- Standardize layout graphics (ie size < 20kbytes, resolution = 72DBI, format=JPEG)
- Consider article organization for archive and future access

---

Thinking about an NT network?

- James W. Brunt, LTER Network Office

I get the opportunity to review many plans and proposals from people jumping on the Microsoft Windows NT; (NT) operating system (OS) bandwagon - oftentimes these plans have not adequately considered how to configure a network that efficiently uses the power of this OS. Typically, these plans propose to use a centralized network model that is typical in UNIX; systems (a large server and a bunch of workstations). This model is not appropriate for an OS like NT that was designed for distributed client/server computing - the model has been successfully propagated because; 1) we are comfortable with it, and 2) it sells better than distributed computing. NT is a true client/server operating system - UNIX on the other hand gives us a taste of

client/server computing but is really a host-based OS. A distributed resource model is much more appropriate to client server computing. This affects the distribution of resources in your network. The biggest mistake in moving from centralized computing to distributed computing is not putting the power on the desktop where the bulk of applications will be (should be) running. The second pitfall is investing in one big powerhouse server - expecting it to behave like a UNIX server. There is an enormous amount of computational power tied up in NT servers out there that is rarely tapped while users machines are desperately stressed.

Having spent over a decade in the UNIX environment, it was essential for me to recognize that there is a paradigm shift that must be acknowledged to tap the power of NT. Creating a working knowledge of the configuration of computers in a distributed client/server environment will help to increase the computing resources available to scientists. There is no justification for the use of a single, large powerhouse server in an NT network especially, if the server will be doing nothing more than basic mass storage, file service, and minor network management. It is almost always better to rely on a group of moderate servers. In true client/ server operating systems like 'NT', it is important to distribute the resources and have redundancy. NT is not a multi-user operating system and is narrowly multi-tasking. A group of moderate servers will easily out-perform one powerhouse server and provide better than adequate file service.

Provide Redundancy - In a managed 'NT' network functionality is 'greatly' reduced if the Primary Domain Controller is not available - it's pretty much a necessity to have a Backup.

Distribute Resources - a good rule of thumb in NT is to have a server for every **major** network function - file service, email service, web service, database server - then add servers as resources become over utilized. Again, servers don't have to be powerhouse machines - for file service machines you need memory and diskspace. Desktop machines need graphics, memory, and processors, and sometimes lots of disk space. Desktop machines in the average 'NT' network should be as or more powerful than the servers. Servers may be more expensive to configure because of power, disk, and peripheral configurations, but in general, the ratio of dollars spent on servers vs. desktops should be close to 1:1 for power application users and 3:1 for administrative users.

Hypothetical Example: If you have $50K budget to set up an NT network to support a 10 -member scientific workgroup. You can buy a $30K server and surround it with 10 $2K desktop machines. The server does email, web service, file service, database service, network management, and license management - this type of system would work fine under UNIX but will not function under NT. Do not mix network intensive applications, processor intensive applications, and data intensive applications. Buy 2 $10K servers and spend more on the desktop machines. One server can do web and database service and the other can do network management, file service, and email - as load increases you may have split some of these functions.

There is a caveat to this model - if you want to get into using NT with your existing system and only have $10K to spend - by all means buy one good server.

Additional Reading: **Microsoft Windows NT; from a Unix; Point of View.** 1995. White paper from the Business Systems Technology Series. Microsoft Corporation.

NACSE-LTER Information Specialists Receive Funds for Database-Spreadsheet Interoperability Project

- Susan G. Stafford, Colorado State University

In 1998, Cherri Pancake, Director of the Northwest Alliance of Computer Science and Engineering (NACSE) and Susan Stafford (LTER DMC Chair) received funding from NSF's Database Activities Program for the OCID (Oregon Coalition of Interdisciplinary Databases) proposal. The thrust of this proposal is to build a set of web-based tools for mining data in collections of interdisciplinary databases. This collaboration has resulted in the refinement of NACSE-developed scripting languages (Hyper-SQL) for dynamic searching (including query refinement) of databases from the Internet. Don Henshaw (AND), Hazel Hammond (AND), and Robin Stubbs (NTL) have cooperated with Joe Hanus (NACSE) to build a demonstration interface to a snapshot of the data in LTER Network climate database, ClimDB. This example interface demonstrates the capabilities of this user interface design tool (HyperSQL) to provide access to this data source. Hazel attended the high-performance networking and computing conference, SC98, contributing to the NACSE research exhibit by showing the CLIMDB interface and discussing implementation issues.The ClimDB interface, developed and maintained by Robin Stubbs of the North Temperate Lakes LTER site, is accessible at http://sql.lternet.edu/climdb/climdb.html.

NACSE and DataTask (DMC steering group) met in February 1998 to discuss potential areas of collaboration. One specific area NACSE has been working on is how to support interoperability between the databases proper and spreadsheet software. NACSE has recently added the capability to control what gets downloaded from a database into a spreadsheet into all 3 of NACSE middleware products - Query Design, HyperSQL, and Query Markup - Language. NACSE is now working on going the other way - being able to upload new data from a spreadsheet and insert it into the database proper. This is not a trivial task because there are some tricky issues like safeguarding that the original database isn't corrupted if the user doesn't have as many fields as the researcher thinks or there are format inconsistencies between the spreadsheet and the database. This was a project that specifically arose out of the NACSE/LTER meeting. As we all have experienced, many ecologists are wedded to the use of spreadsheets and any mechanisms we can employ to extend the usability and future utility of these spreadsheet based datasets will be greatly beneficial to the general community.

We continue to work with NACSE on pursuing additional sources of funding to make it possible to consider more LTER-related projects.

Site Information Manager-Network Office Exchanges

- Karen Baker, Palmer LTER and James Brunt, LTER Network Office

What does an LTER Network Exchange mean for LTER Information Managers? In recognition of the benefits derived from having individual LTER site members directly involved with network activities, a mechanism for short-term fellowship support has been identified within the framework of the LTER Network Office. A part of the goal to improve connectivity and communication is addressed by the ongoing collaboration with

the National Partnership for Advanced Computational Infrastructure (NPACI) which is working to make high-performance environments more generally available. NPACI Funds have been made available to support individuals to work on projects that will accelerate the accessibility of LTER Network data for modeling and synthesis activities. Two pilot projects supported in this manner include ClimDB and SiteDB which are prototype modules of the Network Information System (NIS).

Access to climate data is critical to many cross-site and modeling studies in the LTER Network. The North Temperate Lakes LTER has served as the test development site for the ClimDB prototype through the efforts of Barbara Benson, NTL data manager, and Robin Stubbs, NTL programmer. Robin migrated the ClimDB prototype from the ORACLE database at NTL to Microsoft SQLserver at the Network Office in February of 1999.

Descriptive information about the sites is another critical component of any modeling effort in the LTER Network. The site description database and interface prototype, SiteDB, addresses this need. The Palmer LTER data manager, Karen Baker, is receiving support to further develop SiteDB by exploring network tools available for database descriptions and co-ordinating with GTOS efforts. Starting in 1999, her efforts also include exploration of communication structures such as the network web structure, working document areas, figure archives, electronic newsletter template, and network registration forms as well as establishment of partnerships between information management and education.

These pilot efforts have focused on NIS development activities, but future projects are not limited to NIS work. The Network Office will strive to identify and make available funds for these kinds of activities. Data Managers working on LTER Network oriented projects may inquire about opportunities for support to James Brunt (jbrunt@lternet.edu) or Robert Waide (rwaide@lternet.edu) by outlining a project that would benefit from network support in the form of salary support for a site data manager or programmer, travel of network office personnel to the site and/or travel of site personnel to the Network Office. NIS related projects should continue to be routed through the NIS working group and the Data Management Committee.

◆ *News Bits*

Database Design Tools

- Karen S. Baker, Palmer LTER and James Brunt, LTER Network Office

Database Design Tools are used to help create data models and maintain database applications in conjunction with most of the popular relational database systems (ie Oracle, Sybase, SQLserver, mSQL, mySQL). Data modeling applications permit logical and physical database design and construction making it possible to visualize structure and key elements. Features include bi-directional synchronization of logical and physical designs, reverse-engineering of databases, and HTML-based reporting capabilities. The main drawback today remains the high cost of several thousand dollars.

ERWin (http://www.platinum.com/products/appdev/erwin_ps.htm) is used at the Andrews LTER site to forward and reverse engineer their data model into/from both SQLServer and Foxpro. Much of the SQL code creating tables and triggers is handled by ERwin.

ER/Studio (http://www.gbsweb.com/Database_Tools/ER_Studio/er_studio.html) is used at the LTER Network Office both to design a database as well as to output a structure design diagram of existing databases that can be presented on the web for group viewing.

CAP Web Bird Survey and Catalog

- Peter McCartney, Central Arizona-Phoenix Urban LTER

A data browsing application has been written for a bird survey project. Be sure to look at the transect map portion that was written using Visual Basic and ESRI MapObjects. Data queries are done using Active Server Pages. We expect to develop similar presentations for most of our monitoring projects (http://caplter.asu.edu/po12).

A CAP LTER searchable data catalog is online (http://caplter.asu.edu/datacatalog). There is still much work to be done. Although much of the metadata is incomplete, the online catalog effort is started.

Electronic Multi-Authoring

- Karen S. Baker, Palmer LTER

How can a co-author edit a multi-authored paper electronically and still provide the explicit visual context of a marked up manuscript for each change or addition? LTER Data Managers have begun using marked up electronic text to pass suggested modifications to a primary posting author for integration into the final document. Text modification can be made locally to a downloaded document using readily available software.

A manuscript may be posted at a web location (URL) as an HTML, RTF, or Word document. The document may be downloaded (using FTP, Netscape or Internet Explorer) and stored locally. The downloaded file can be viewed with an editor such as Microsoft Word. The 'track changes' in this editor means that the deleted text can appear as colored (ie blue) with strike outs while inserted text can appear as an alternate color (ie red). These edits can be saved as an HTML, RTF or Word file and posted to a local URL or attached to an email. The original author is able to download or receive the marked up manuscript with edits visible in the Microsoft Word editor. Note that to be visible directly with a browser, highlighting requires html commands font color=#FF0000,#FF3333 substituted for editor commands #7B0516,#7D031E,#7D0322.

Network Updates

- Richard Dahringer and Colin Johnson, LTER Network Office

Microsoft SQLserver LTER databases at the Network Office are currently queried using PERL running under IIS Microsoft Information Server on the NT. Tests are being run using queries initiated from UNIX to NT using a C Library (http://metalab.unc.edu/freetds) of code ported from Linux to Solaris. This permits data processing to take place on the SUN instead of the PC giving a potential improvement in performance.

Investigations are ongoing with the public domain programming language PHP (http://www.php.net) for use with web page communications such as queries and updates to relational databases (ie, mSQL, mySQL, postgresSQL, Oracle, Sybase, SQLserver), XML parsing, as well as graphic generation on the fly.

The lternet.edu web revision plans to improve the web page load rate are in the planning stages given that the current graphics refresh impacts performance significantly.

Software Bits

Communications software under consideration:
*NetMeeting (http://www.microsoft.com/netmeeting/) is a Microsoft internet shared-screen conferencing tool which allows remote PC users to talk and share electronically by viewing common typing areas and white boards.
*Virtual Network Computing (VNC) (http://www.uk.research.att.com/vnc/) is free software that creates a remote display system allowing viewing of a computing environment on any machine, ie a unix desktop from a pc or a pc desktop from a unix platform, which means a window with your local envirnoment can be opened on another machine. (C.Johnson)
* StarOffice (http://www.stardivision.com) provides a solution to the mail reading for those people running UNIX and wanting to read mail generated on other platforms. StarOffice permits those running unix to read Microsoft word files. A "free for non commercial use" written in Java so availabe for Windows, OS2, Linux and Solaris includes a MS word like word processor as well as Excel, Powerpoint and Drawing "like" programs (D.Blodgett).

◆ *Calendar*

*99 Jan 18-22 Network Office Visit by K.Baker (PAL)
*99 Feb 08-09 LTER Executive Committee meeting, Washington, DC

*99 Feb 00-00 Network Office Visit-ClimDB by M.Stubbs (NTL)
*99 Feb 18-23 Data Manager DataTask Meetings, NCEAS
*99 Feb 23-25 Network Visit-SiteDB by D.Blodgett (BNZ) & K.Baker (PAL)
*99 Mar 25-28 LTER Technology Committee Meeting, SDSC, California
*99 Apr 22-23 LTER Coordinating Committee meeting, Luquillo LTER, Puerto Rico
*99 Jun 22-25 Workshop on the Central Euopean LTER Network, Budapest, Hungary
*99 Aug 04-06 LTER Data Manager Committee Meeting, Spokane, Washington
*99 Aug 07 ESA/LTER Workshops, Spokane, Washington
*99 Aug 08-12 ESA Annual Meeting "Landscapes, Legacies & Limits Bridging Borders" Spokane
*99 Aug 15-21 Annual Meeting International LTER Network, Kruger National Park, South Africa
*99 Oct 01-03 LTER Coordinating Committee, Hubbard Brook LTER (science theme: Biodiversity)

LTER Education Workshop 22-24October98

- Diane Ebert-May, Chair and Patty Sprott, Logistics

The LTER Education Committee (http://www.lternet.edu/oppts/education/) held the first LTER Education Workshop 22-24 October 1998 at Biosphere 2 in Oracle, Arizona. Funded by NSF (Environmental Biology and Education & Human Resources), each of thirteen LTER sites sent a three member team (a scientist, an educator, and a teacher). The focus was to develop approaches to education at the LTER sites. Two LTER site Data Managers (Hap Garritt/PIE and Karen Baker/PAL) took part in the developing Science-Education-Information Management Partnership.

LTER DataTask Meeting 18-23Feb99

- Susan Stafford, Chair

A DataTask and NIS subgroup Meeting was held at NCEAS in Santa Barbara with Chris Wasser, Robin Stubbs, John Porter, Karen Baker, Darrell Blodgett, Peter McCartney, Don Henshaw, James Brunt and Susan Stafford attending. We were joined by Matt Jones (NCEAS), Bob Waide (NET), Tony Fountain (SDSC-LTER postdoc) and our old friend Rudolf Nottrott (NCEAS). We began the meeting with the major objective of continuing our discussion of the NIS but after Bob Waide's report and his charge to our committee to develop a White Paper on "Future Initiatives, Directions, and Challenges for LTER Information Management", we quickly folded our NIS effort into the context of this broader planning effort. We discussed the creation of a Bio-Informatics Consortium (BIC) in which LTER would be a major player.

LTER Technology Committee Meeting 26-27Mar99

- John Vande Castle, Chair; Tony Fountain, Logistics

The LTER Technology Committee (http://lternet.edu/research/technology/) met at SDSC with seven data managers among the participants (John Anderson/JRN, Karen Baker/PAL, Emery Boose/HRF, Hap Garritt/ PLI, Jim Laundre/ARC, Gregg MacKeigan/SEV, Sam Walker/BES, Chris Wasser/SGS). Speakers gave overviews of regional modeling (Stuart Gage/KBS), remote sensing (Greg Asner/NWT) and data management (Sam Walker/BES) prior to break out groups meeting to discuss future LTER technology needs in the realms of communications, biotic and physical/chemical measurements.

LTER Data Manager Meeting 04-06Aug99

- Susan Stafford, Chair; Chris Wasser and Barbara Benson, Coordination

The full agenda of DataTask indicates a busy August agenda in Spokane so we are planning to make information available prior to our upcoming summer Data Managers Meeting. With an expanding network and more interest in our efforts by the greater community, it will be imperative that participants come prepared and the charge to working groups be defined ahead of time. All leaders of meeting sections are asked to come with their presentations " in hand" and posted on the meeting web page. A draft list of discussion topics includes:

- Year 2000 meeting (Wasser, Benson)
- LTER IM Vision White Paper (Stafford)
- NIS (Henshaw)
- New Federal Opportunities and KDI proposal (Waide/Brunt)
- Network Update (Brunt)
- Education/IM interface (Baker)
- Sociology/Ecology interface (McCartney)
- Information for new sites (Hartman)

DataTask requests that additional topics of interest be forwarded prior to summer. Note ESA workshops begin 07 Aug and ESA meetings 08-12 Aug (http://www.sdsc.edu/esa/99meet.htm).

01001100 01010100 01000101 01010010