All these access methods are provided but without elaborate page structure. Duplication under each access method, if it serves the reader, is allowed. Since screen real estate is at a premium and short load time is very desirable (this is, after all, a list for those *at work.*), simplicity is the rule: the structure is that of a simple list indented as appropriate.

Effect of Drivers

Fermilab top level page and immediate subsidiary pages:

The effects were:

1. The home page tended to be more static and less easy to change.
2. There were more severe strictures on form and content of the top level and second level pages.
3. There was more sign-off required to make a change to any of these pages.

These were intentional and desired. A technology which achieves prominence as a communications channel to a large general audience needs to be managed carefully and in a fashion that is consistent with the organization's overall direction. In addition, as people have come to depend upon the pages for information as part of their regular everyday activities, stability and familiarity have become important–particularly at the top page levels.

Collaboration and Support Group pages:

The main effect was that collaborations and support groups were asked to designate a "local webmaster" as a prerequisite to having their web server listed as a link off the Fermilab home page (e.g., in the Fermilab Experiments & Projects portion of the Fermilab At Work [HREF35] page).

This "local webmaster" is expected to be a contact with whom the Fermilab webmaster could discuss technical issues (e.g., server down, change of server node, etc.). The local webmaster is also expected to make sure that the server content is consistent with various proper computing usage policies. A distinction is made between collaboration/support-group "institutional" pages and home pages for individuals. The "local webmaster" is held responsible for the content of the "institutional" pages. The individuals are held to conformance with proper computing usage policies in the content of their home pages. When setting up the individual's home page, the local webmaster is expected to remind individuals about these policies which they have seen before at the time they received their computer account.

An element of "line management responsibility" was put in place by requiring the local webmaster's supervisor (or official spokesperson in the case of collaborations) sign a form designating the particular individual as local webmaster. In addition, a supervisor is strongly encouraged to obtain the blessing of their Division or Section head in order to make sure there is a clear understanding at a high level about what is being done.

Consequences:

With respect to the top level pages, the consequences have been favorable. The structure has held up over the year since the Fermilab Home Page was created and seems to have served its purposes well.

With respect to the collaboration/support group pages and servers, there has been a modest amount of comment about the sign-off bureaucracy. This has been weathered. The delegation of responsibility for content to an accountable, recognizable line management chain has given management (and certainly the webmaster) a certain comfort level in linking independent servers to the Fermilab Home Page.

At least in one circumstance, the arrangement worked well when a particular content was questioned. Thus, the procedure appears to work and, most importantly, the collaborative use of the web and web servers at Fermilab appears to be healthy, on-going, and adopted in greater and greater numbers.

Library Web Architecture

Overview:

With a staff of six the Fermilab Library is responsible for the information and research needs of some 700 scientists and another 2,000 laboratory employees. There is a collection of 15,000 book volumes, subscriptions to 250 journals, and a weekly receipt of approximately 200 preprints. Preprints, research papers which are circulated among the high energy physics community worldwide, are the most important part of the Library's collection. Preprints in paper format are diminishing. They were one of the first sources to move to full-text online format.

When a library's major collection changes formats, focus on access must shift. The typical Fermilab Library reader is highly literate in electronic media and in turn has high expectations in regard to information access and delivery. At the same time preprints were migrating to such places as the xxx e-Print archive [HREF36] at Los Alamos National Laboratory [HREF37] and to any of a number of other preprint bulletin boards [HREF38] as well as onto WWW servers of specific research individuals and groups, more and more information sources were also becoming available via the web. Through its own single web server [HREF39], the Library was able to direct its readers not only to WWW resources, but also to gopher, ftp and telnet sites. WWW has become a common denominator in this high energy physics setting as a way of making active references to information sources.

Goals and their Implications

Immediacy of access to information resources:

Since many documents of interest to the high energy physics community are now being generated on web servers, web access is often the quickest and most efficient means of retrieval. The choice of a web server over a gopher server or ftp access was based on the following:

1. the natural affinity between the web and high energy physics research
2. the ease of server implementation
3. the growing web infrastructure in the Laboratory
4. the flexibility of the URL structure allowing for connections not only to other WWW servers, but also to gopher, telnet, ftp sites.

Relevancy, both in content and format, to reader and laboratory information needs:

In order to achieve and maintain relevancy to reader needs the Library's web offerings are fluid in design and change according to reader demand, both actual and anticipated.

With its own server, the Library is free to make links, change links, break links, and change the design of the hierarchy of resources–all in conjunction with the ever-changing reader needs. As the maintainer of the Library's web server, the head librarian incorporates new web resources discovered by means of listservs, library literature, library and laboratory colleagues. Most of these are placed under the "Hot Links" [HREF40] section of the Library's home page.

Centralization and marketing of Library resources:

The Library home page brings access to diverse resources together in one place. It is a "clickable" library brochure which, moreover, can always be current. Since automating the Library's catalog in 1990, the idea had been to customize the automation system to be the platform for jumping off to other databases and resources.[8] The Library home page has in fact become that "jumping off" point. The automation system, the principal repository of bibliographic and holdings information, is now coming to be thought of as just another information source among those provided by the Library. The web is what is bringing them all together.

Empowerment of the Library staff to author their own web information sources:

The Library server was installed on a Macintosh already in place in the Library. This was familiar territory for the librarian which increased the probability of the server's success. In addition, the experience the librarian gained in writing HTML documents and navigating the web made it possible for the Library to become one of the organizations in the Laboratory that could provide information, guidance, and mentoring on WWW, HTML, and URL's. Most of the Library's web offerings are links to other servers rather than original documents, but this of course reflects the referral nature of library services.

Content and Structural Drivers

Reader Interest:

Distinctive from the Laboratory web server, the factors driving the content and structure of the Library server are much more closely tied to reader interests and needs. Setting up of the Library web server in April 1994 was motivated by reader need to gain access to two preprints which presented evidence of the existence of the top quark, the (then) last, remaining, yet to be discovered quark in the Standard Model theory linking all sub-atomic constituents.

Links to the full-text of these two preprints were the first ones the librarian added to the Library server's "Hot Links." Also handling most of the reference queries, the librarian increasingly turned to WWW first in tracking down government documents such as Presidential press releases and science policy statements requested by the Laboratory's Director and other readers. Good URL "finds" generated from this work were often times incorporated into the Library's "Hot Links" [HREF41] section.

Organizing the web:

When found on the web, links were quickly incorporated into the Library's server often "bumping down" or replacing older links. For example, links to the "top quark evidence" preprints were moved down the Hot Links list as reader interest waned. Now those first links have been removed altogether and replaced by links to the top quark discovery papers [HREF42] by the CDF and D0 collaborations. Again, when significant documents concerning governmental support of science were released, links were added. Two examples here are the U.S. White House report, "Science in the National Interest," [HREF43] and the Drell Report [HREF44] issued by the U.S. Department of Energy on the future of high energy physics. As these documents became less topical, links were removed.

The frequency of the changes is an element of the chosen architecture for the Library's web server: reader-driven change.

Hardware/Software Availability:

As previously noted, the Library server runs on existing Mac hardware. The MacHTTP server and MacMosaic browser software was downloaded from the National Center for Supercomputing Applications [HREF45] (NCSA) in Urbana, Illinois. The easy installation obviated lengthy searches for staff resources that would have been necessary had the software been more difficult to install.

Implications for Libraries and Library Administrators

Involvement of the Library Staff in the Web Architecture

Make Sure there is an Understanding of Basic Concepts:

With a conceptual understanding of the fundamentals of web architecture, library staff can build the confidence needed to maximize their use and in turn the readers' use of web sources. Involvement in the initial period of setting up web server and browser, as well as many other technological tools, provides a sense of "owning" that technology which often motivates future participation in such projects.

Encourage Interaction between Librarian Staff and Computing Professionals:

In the spirit of collaboration, interaction between library staff and computer professionals can foster exchange of ideas and mutual professional respect. The potential for a successful web service is heightened when technology is combined with understandings of usage patterns and needed content. The library staff know their web users–their readers–and their ways of getting from one topic to another. The library staff need to participate in designing the links and other architecture for the web server so that readers will find the linkages natural.

Support the Provision of Continuous Training and Access to Developments in Technology:

Support from library administrators of continuous training programs for library staff along with the provision of hardware, software, network upgrades, etc., is imperative. Library staff will have to have the tools, the training, and the time to experiment with expanding technology and electronic information sources.

New Face of Library Services

Library services have a new face brought on by the new technological capabilities. The pace of development is so fast that one must plan while implementing and implement while planning.

The focus in this paper has been Reference Services but changes are occuring in other service areas as well. It is entirely possible to imagine multiple web servers within a large library–some devoted to circulation (e.g., to support reader access to his/her own circulation records), technical services (e.g., to provide information on recently cataloged materials), and reference (e.g., to provide access to "Hot Topics" pages maintained by individual reference specialists for their particular specialty fields). Indeed, a number of libraries already are making innovative uses of web technology along the lines described.

See for example, Innovative Internet Applications in Libraries [HREF46] maintained by Ken Middleton of Todd Library, Middle Tennessee State University and also Fisher Library at the University of Sydney [HREF47].

When a library begins to have multiple servers with multiple collections of pages maintained independently by varied individuals across the institution, the web organization issues begin to resemble those for a large laboratory, such as Fermilab. It may be that the solutions described here will also apply.

New Role for Libraries within Institutions and Communities

The work of librarians as information specialists has suddenly become popular within our culture. However, the general populace does not necessarily associate librarians with the rise of electronic information. The need for consciousness raising about the library profession has never been greater. The doors are open to librarians to step in as their institution's or community's mentors for Internet access and tools such as WWW.

The ease of gathering and authoring via the web empowers librarians with a greater independence and flexibility. As a result, they will be able to change their library's public interface at a pace more in step with reader needs instead of being so highly dependent on vendors, computer professionals, or others.

The Next Phase of Library Automation

With web technology, systems centered around a single vendor or single information provider give way to a reader-centered collection of many systems and information resources accessible over the network. The unique aspect about the web development has been the *ease* of access along with the *quickness* of switching from one information source to another.

The web's "Uniform Resource Locator" (URL) concept is key to making this ability to switch quickly possible as is, of course, the underlying presence of that uniformly accessible, totally interconnected, network of networks, which is the Internet. Together, the two provide a seamless interface to these varied information resources, bringing on the next phase of library automation by creating what one might call the information marketplace.

Of course, the information marketplace has existed for a long time–the pivotal aspect of this new information marketplace is that the point of purchase has changed. Formerly, the decision to go with this or that commerical database vendor was likely made in the library administrator's office once a year at contract time. Now, the decision is made by the individual reader as he or she chooses this or that network information resource on possibly a minute to minute basis.

Currently, in the "spirit of the internet," there are many "vendors" supplying information for free. Given the ease of authoring and making documents available on the web, this may continue for some time. Still, the demand for authoritative documents or those with other special characteristics accessible in the same way as the free material is giving rise to vendors who provide access to information resources on the network for a price.

Vendors naturally work to have their resources made essential and placed at the center. Special browsers, special servers, or other software that use the underlying web and network transmission rules are all possible. Regardless of what vendors devise for providing access to their information over and above the basic Mosaic browser kinds of capabilities, libraries must require conformance to standards and mandate interoperability (e.g., require that information servers from various vendors work with a wide selection of information browsers).

Not only will this force competition into the information marketplace, encouraging high quality and low price, but it will also allow the library to stay "light on its feet" and very adaptable to the highly diverse, growing numbers of information resources on the Internet.

By so doing, libraries will insure that they, in the next phase of automation (which in fact is now occuring), can continue to be reader-centered rather than system-centered and thereby be of maximum benefit to their reader communities.

Broader Implications

The web architecture within an organization is shaped by many different drivers. The institution's top-level web server is likely to be very structured while the library's web structure, particularly in the reference area, needs to be flexible, dynamic, and responsive to reader interest. The democratization of information authoring and access which the web supports makes it possible to bring many more people into the process of information distribution and access.

Conclusion

The advent of World Wide Web has brought together diverse, formerly distant parts of the Laboratory. Suddenly. press officers, publications specialists, patent lawyers, librarians, and researchers are finding the focus of their work to be just a mouse click away from one another. In collaboration, the Library and the Laboratory have devised a web architecture which makes information accessible to readers while respecting the policies of funding agencies, the sensibilities of authors, and the concerns of other information stakeholders.

References

T. Berners-Lee, R. Cailliau, "World Wide Web," *Proceedings of 10th Conference Computing in High Energy Physics*, Annecy, France, September 21-September 25 1992, 69.

T. Berners-Lee, R. Cailliau, A. Luotonen, H. Frystyk Nielsen, A. Secret, "The World-Wide Web," *Communications of the ACM* 37, (1994) 76.

Hypertext References

HREF1	http://www.fnal.gov - Fermilab
HREF2	http://www.w3.org/hypertext/WWW/TheProject - WWW
HREF3	http://www.cern.ch - CERN
HREF4	http://fnnews.fnal.gov/cd/ - Computing Division
HREF5	Online Systems Department - http://www-ols.fnal.gov:8000
HREF6	http://fnala.fnal.gov:8000/docdb - Server to provide online data acquisition system documents
HREF7	http://www.w3.org/hypertext/WWW/People/Berners-Lee-Bio.html - Tim Berners-Lee
HREF8	http://www.ncsa.uiuc.edu/ - NCSA
HREF9	http://www.ncsa.uiuc.edu/SDG/Software/Brochure/MacSoftDesc.html#MacMosaic - NCSA MacMosaic
HREF10	http://www.ncsa.uiuc.edu/SDG/Software/Brochure/PCSoftDesc.html#WinMosaicNCSA - WinMosaic
HREF11	http://www.ncsa.uiuc.edu/SDG/Software/Brochure/UNIXSoftDesc.html#XMosaic - NCSA XMosaic
HREF12	http://fn781a.fnal.gov - E781
HREF13	http://www-sdss.fnal.gov:8000/ - The Sloan Digital Sky Survey

HREF14 http://www.fnal.gov - Fermilab

HREF15 http://www.fnal.gov/top_news_release94.html - Announcement of evidence for the top quark

HREF16 http://www.fnal.gov/top95/top_news_release.html - Announcement of evidence for the top quark

HREF17 http://www-cdf.fnal.gov - CDF

HREF18 http://d0sgi0.fnal.gov - D0

HREF19 http://libmc1.fnal.gov - Fermilab Library web server

HREF20 http://www-sdss.fnal.gov:8000/ - The Sloan Digital Sky Survey

HREF21 http://www-dart.fnal.gov:8000/ - DART

HREF22 http://www.fnal.gov/hep_descript.html - Science of High Energy Physics

HREF23 http://www.fnal.gov/fermilab_intro.html - Laboratory as a Research Institution

HREF24 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF25 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF26 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF27 http://www.hr.doe.gov/doevideo/doedtsad.html - U. S. Department of Energy Televideo Service

HREF28 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF29 http://www.doe.gov - U. S. Department of Energy

HREF30 http://www.whitehouse.gov - The White House

HREF31 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF32 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF33 http://www-fermitools.fnal.gov - Fermilab Software Tools Program

HREF34 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF35 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF36 http://xxx.lanl.gov/xxx - e-Print archive

HREF37 http://www.lanl.gov/ - Los Alamos National Laboratory

HREF38 http://libmc1.fnal.gov/Hot.html - relevant preprint bulletin boards

HREF39 http://libmc1.fnal.gov - the Fermilab Library's web server

HREF40 http://libmc1.fnal.gov/Hot.html - "Hot Links" on the Fermilab Library web server

HREF41 http://libmc1.fnal.gov/Hot.html - "Hot Links" on the Fermilab Library web server

HREF42 http://www.fnal.gov/top95/top_discovery.html - links to the top quark discovery papers

HREF43 http://docs.whitehouse.gov/white-house-publications/1994/08/1994-08-03-science-in-the-national-interest-policy-statement.text - "Science in the National Interest"

HREF44 http://www.hep.net/documents/drell/full_report.html - Drell Report

HREF45 http://www.ncsa.uiuc.edu - National Center for Supercomputing Applications

HREF46 http://frank.mtsu.edu/~kmiddlet/libweb/innovate.html - Innovative Internet Applications in Libraries [HREF46]

HREF47 http://www.library.usyd.edu.au - Fisher Library at the University of Sydney

Footnotes

* Note: Double underlined words in the paper version of the document indicate the hypertext references in the network-resident version of the paper submitted to the conference. For the network-resident version, see the URL: http://www.scu.edu.au/ausweb95/papers/libraries/garrett/

For information more generally about the conference, see the URL: http://www.scu.edu.au/ausweb95/

This document was prepared with MicroSoft Word and converted to HTML for the network-resident version using rtftohtml. For further information, see ftp://ftp.cray.com/src/WWWstuff/RTF/rtftohtml_overview.html.

1 Head Librarian, Fermi National Accelerator Laboratory, January 1987 - March 1995.

2 Webmaster, Fermi National Accelerator Laboratory.

3 T. Berners-Lee, Presentation given at Fermilab, July 1992 (unpublished).

4 J. Streets, private communication.

5 Penelope Constanta-Fanourakis, Guide to Using the Document Database - DOCDB, PN 336, March 29, 1988, (unpublished).

6 The top quark is the last to be discovered quark in the Standard Model theory linking all sub-atomic constituents. Researchers have searched for it in physics experiments for some eighteen years. In April 1994, evidence for it was announced by the Fermilab CDF and D0 Collaborations. In March 1995, CDF and D0 announced its discovery.

7 For example, in March 1993, one experiment's private unpublished meeting minutes requested collaborators to store material in the experiment's VAX Notes electronic "conference," a vendor-specific, platform-specific, system somewhat like a bulletin board. By June 1993, the experiment's minutes had announced V1.0 of the experiment's "Documentation System," a WWW Server and implied the demise of VAX Notes. The sense in the announcement was that merely of setting up another channel of collaboration communication--of little relevance to the laboratory Directorate and the laboratory's public posture.

8 P. Garrett and D. Ritichie, "Fermilab Library Projects," FERMILAB-TM-1667 (1990) and P. Garrett and D. Ritchie, "Fermilab Library Directions," FERMILAB-TM-1668 (1990).

Ka-Kung Yee ... Mathematical ... for the ... 01/21/1999 [...]

Stanford Linear Accelerator Center

$e^{i\pi} + 1 = 0$
HTML Math

The SLAC Web Users Group Presents...

Ka-Ping Yee "A Concise Mathematical Expression Language for the Web"

Date:	Tuesday, January 26, 1999
Time:	2:00 pm
Location:	Orange Room
Speaker:	Ka-Ping Yee
Topic:	Math and the World Wide Web

The ability to have print-quality rendering of mathematical expressions on the Web has been a goal since the early days of the technology. Why has realizing this goal been so problematic and what is the status of proposed solutions?

The presentation will demonstrate MINSE, a system created by the author in 1996 for the representation and display of mathematical expressions within HTML documents. The author will discuss its design goals, which include convenient direct entry and accessibility from existing Web browsers, and how these goals were achieved in the language and the implementation.

Ka-Ping Yee (better known as !Ping) has been involved with the Math Working Group (MathWG) of the World Wide Web Consortium (W3C). He currently works as a systems developer at Industrial Light and Magic.

Page owner: *Henniss* .
Last Modified Jan 21 22:56:47 1999 by Henniss.

history, 4/2000

Saving early Web pages before they're history

By Jeanene Harlick
STAFF WRITER

Old Web pages don't die, they just fade away.

In fact, Web pages are blinking away so fast that the very history of these pioneer Internet years is threatened with extinction.

That's why a small group of scientists and students is racing to find ways to archive Web pages, just as paper documents like the original Declaration of Independence have been preserved.

The difference, said archivist Jean Deken, is that paper persists and, so far, objects in the digital world don't.

"They don't give any warning they're going to corrupt. They simply very quietly pass out of usefulness," said Deken, who is with the Stanford Linear Accelerator Center.

A few years ago she found herself trying to electronically chase down the nation's first Web page, born in 1991 in a SLAC physicist's office.

No record of the page existed. Plowing through old backup tapes, the center's Web master found it, a page of plain text that made virtual history.

If the pair had waited one year, the page would have been lost forever, trapped in a file created by obsolete software no current hard drive can read.

It's too late, however, for the first e-mail ever sent. Too late, as well, for countless original World Wide Web pages and other documents stored electronically.

Confronted with voluminous electronic records, storage media that is obsolete within a few years and employees who delete documents before anyone realizes they're important, some fear future historians will look back on this time as a "digital dark ages."

"The chronicle of our entire period (is) in jeopardy," computer scientist Jeff Rothenberg wrote in a 1995 article that sent a shock wave through the digital community. "It is only slightly facetious to say that digital information lasts forever — or five years, whichever comes faster."

Information overload

A look at the National Archives and Records Administra-

vancements to manage them," Carlin told Congress in October. "Unless we successfully address the key issues, essential evidence will be lost."

Lost in the ether

The first challenge posed by digital documents is one of physical presence. Many exist only in electronic format. They are e-mails and Web pages deleted and revised before a permanent copy is retained.

"These really are incredibly fugitive archives," Deken said. "They disappear. They disappear unless somebody takes appropriate steps early enough."

Archiving has been particularly hard for the Linear Accelerator Center because the center records most of its work on internal Web servers accessible to physicists worldwide. Researchers often update individual Web pages without saving the first version.

Because the Linear Accelerator Center is funded by the Department of Energy, it is required to send National Archives documentation of big projects like GLAST, the gamma-ray telescope the lab is now building.

Currently, Deken and Web master Joan Winters deal with the problem by sifting through backup tapes — which take periodic snapshots of the center's Web pages — every year to find important records. It's a time-consuming and inefficient process, and one Deken hopes will soon change.

Deken said organizations like the Linear Accelerator Center need a computer program that somehow automatically archives important Web documents as soon as they're created.

Change, change, change

The second challenge is the instability of electronic media.

Floppy disks are easily

Stanford Linear Accelerator Center Operations Manager John Weisskopf peruses storage racks of computer files Monday afternoon.

BRET PUTNAM — Staff

erased by stray magnetic fields or simple material decay.

But the more vexing problem is that hardware used to create documents and software used to interpret them are obsolete within five years, according to a National Media Lab study. Archivists are left with a disk — or CD, or whatever future records are stored on — full of ones and zeros as impossible to translate as hieroglyphics before the Rosetta Stone.

For example, eight-inch floppy disks — the standard storage media for e-files in the 1970s — are now unreadable.

It's like trying to find a place to play that old eight-track of Willie Nelson's greatest hits.

At the Linear Accelerator Center, they created their first Web pages using an operating system called IBM Virtual Machine. Five years later, when Winters went looking for the pages, VM was being phased out and replaced by Unix.

If Winters had waited until October 1998, when the center's VM machines were chucked, the pages would have been trapped in a virtual no man's land.

What's the solution

The Internet Archive, a nonprofit organization in San Francisco, started recording the World Wide Web in 1996.

Using Web crawlers that automatically collect pages from public servers, the organization gets a complete snapshot of the Web every two months.

They don't give any warning they're going to corrupt. They simply very quietly pass out of usefulness," said Deken, who is with the Stanford Linear Accelerator Center.

A few years ago she found herself trying to electronically chase down the nation's first Web page, born in 1991 in a SLAC physicist's office.

No record of the page existed. Plowing through old backup tapes, the center's Web master found it, a page of plain text that made virtual history.

If the pair had waited one year, the page would have been lost forever, trapped in a file created by obsolete software no current hard drive can read.

It's too late, however, for the first e-mail ever sent. Too late, as well, for countless original World Wide Web pages and other documents stored electronically.

Confronted with voluminous electronic records, storage media that is obsolete within a few years and employees who delete documents before anyone realizes they're important, some fear future historians will look back on this time as a "digital dark ages."

"The chronicle of our entire period [is] in jeopardy," computer scientist Jeff Rothenberg wrote in a 1995 article that sent a shock wave through the digital community. "It is only slightly facetious to say that digital information lasts forever — or five years, whichever comes faster."

Information overload

A look at the National Archives and Records Administration shows just one example of the depth of the problem.

During the past quarter-century, the National Archives has been able to process just 100,000 e-records. But the Treasury Department alone is generating nearly one million e-mail messages a year that should be permanently preserved, U.S. Archivist John Carlin estimates.

In 1976, when the National Archives and Records Administration went looking for the 1960 U.S. Census data, officials realized it was stored on UNIVAC tapes whose tape drives — the first main frame computers — no longer existed. It took three years and some engineering back flips to eventually recover the data.

In addition, the National Archives has no digital archives or guidelines on how federal agencies should store electronic records the agencies are mandated to preserve.

"New technological ways to create records have not been matched by technological ad-

[illegible] erator Center because the center records most of its work on internal Web servers accessible to physicists worldwide. Researchers often update individual Web pages without saving the first version.

Because the Linear Accelerator Center is funded by the Department of Energy, it is required to send National Archives documentation of big projects like GLAST, the gamma-ray telescope the lab is now building.

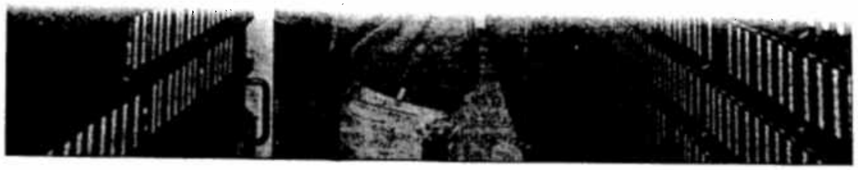
Currently, Deken and Web master Joan Winters deal with the problem by sifting through backup tapes — which take periodic snapshots of the center's Web pages — every year to find important records. It's a time-consuming and inefficient process, and one Deken hopes will soon change.

Deken said organizations like the Linear Accelerator Center need a computer program that somehow automatically archives important Web documents as soon as they're created.

Change, change, change

The second challenge is the instability of electronic media.

Floppy disks are easily



erased by stray magnetic fields or simple material decay.

But the more vexing problem is that hardware used to create documents and software used to interpret them are obsolete within five years, according to a National Media Lab study. Archivists are left with a disk — or CD, or whatever future records are stored on — full of ones and zeros as impossible to translate as hieroglyphics before the Rosetta Stone.

For example, eight-inch floppy disks — the standard storage media for e-files in the 1970s — are now unreadable.

It's like trying to find a place to play that old eight-track of Willie Nelson's greatest hits.

At the Linear Accelerator Center, they created their first Web pages using an operating system called IBM Virtual Machine. Five years later, when Winters went looking for the pages, VM was being phased out and replaced by Unix.

If Winters had waited until October 1998, when the center's VM machines were chucked, the pages would have been trapped in a virtual no man's land.

What's the solution

The Internet Archive, a non-profit organization in San Francisco, started recording the World Wide Web in 1996.

Using Web crawlers that automatically collect pages from public servers, the organization gets a complete snapshot of the Web every two months.

f i ® s t m ¤ ñ d @ ¥
PEER-REVIEWED JOURNAL ON THE INTERNET

FM Interviews

Louise Addis

I was so excited that I almost dropped the book.

I was reading *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor,* by Tim Berners-Lee. The paragraph that caused such a reaction described how Paul Kunz, a California physicist, had returned from a visit to CERN (Conseil Européen pour la Recherche Nucléaire; now officially named the Organisation européenne pour la recherche nucléaire or European Organization for Nuclear Research) in Switzerland where he had seen an interesting technology developed by Berners-Lee. Kunz came home to the Stanford Linear Accelerator Center (SLAC) and told Louise Addis, a SLAC librarian, about this new invention, the World Wide Web.

Berners-Lee states that Louise saw the Web "as a godsend - a way to make SLAC's substantial internal catalogue of online documents available to physicists worldwide." [1] He further states that "under Louise's encouragement SLAC started the first Web server outside of Europe." [2]

The first Web server created in the United States - set up anywhere outside of its birthplace at CERN - was established at the encouragement of a librarian!

As someone working in both the Web world and the library community, I decided then that I had to meet, or at least talk to, Louise Addis. I knew that I could learn something from the librarian who saw the potential in this revolutionary technology before the majority of the world had even heard of the Internet.

I tracked Louise down using the Web (of course!) and e-mail. When we first spoke on the phone, I told her how excited I was to meet the person who was responsible for setting up the first Web server in the U.S. Louise laughed and disclaimed total responsibility for bringing the Web to America. She told me about her colleagues, Paul Kunz and Tony Johnson and others, who were critical to the launch of the first Web server at SLAC. At one point in the conversation, she noted that the history of the Web technology is rather like *Rashomon,* Akira Kurosawa's 1951 movie in which the same story is told from widely diverging perspectives. It appears to be true that there are a number of stories about how the Web came to America. This is one of them. - Melissa Henderson.

First Monday (FM): Could you first tell me a bit about your academic background? Where did you study for your undergraduate degree?

Louise Addis (LA): I studied International Relations at Stanford University. I always had an interest in the sciences, however.

FM: And from Stanford?

LA: I initially went to work for a publisher of the "Dick and Jane series" in the claims department and as a sales correspondent. I then moved to the Stanford Research Institute Radio Science Lab where part of my job was handling classified documents. I was thinking about going to library school when a colleague encouraged me to apply for a position in the soon-to-be-opened library for Project M.

FM: Project M?

LA: "Project Monster" now known as the Stanford Linear Accelerator Center. At that time, it was still on the drawing board and had no funding. We started our library in a warehouse on the Stanford campus. It was the hottest place on campus; they ran water on roof to make it more bearable!

FM: How long were you at SLAC?

LA: From before the beginning! It was very exciting to build a library. I started as the assistant to the librarian in 1962 and retired in 1994 as Associate Head Librarian. Since retirement, I have continued to work off and on for the SPIRES-HEP database project at SLAC and most recently on Y2K conversion.

FM: What were you working on in your early years at SLAC?

LA: From about 1969 on, I was working with the SPIRES-HEP database of particle physics documents. SPIRES (initially, Stanford Physics Information Retrieval System, then Stanford Public Information Retrieval System) database management system was the product of brilliant programmers at Stanford. SLAC was a guinea pig for SPIRES, which was originally designed for an IBM mainframe environment. SLAC sponsored development of the Unix version, which is now in use.

FM: Did anything in your education or background prepare you for the higher level computing work you were doing with SPIRES-HEP or the Web?

LA: I got into programming and database development by picking up information on my own. If I needed to know something, I asked someone to show me how to do a particular task. Then I went back to the Library and tried it on my own.

When the SLAC Library was started, the Stanford University computing center was a little building with keypunch machines and a line printer. At the Stanford Research Institute, I had seen an "advanced serials handling system" which meant that each serial had its own punch card. This was cutting edge! So when we built the Library at SLAC, we put the serials on punch cards. I remember walking across campus with my stack of punch cards. So advanced!

FM: What was so important about SPIRES-HEP database of particle physics documents?

LA: There weren't any library database management tools available at the time and SLAC needed something to meet the unique needs of particle physics researchers. Research papers produced by the particle physics community often have hundreds of authors. The record is 1,200 authors! We were committed to listing all authors or researchers and needed a tool that would accommodate us.

Research papers produced by the particle physics community often have hundreds of authors. The record is 1,200 authors.

The SPIRES-HEP database covers the full literature of particle physics and is a collaborative project between a number of particle physics labs, primarily SLAC and DESY in Hamburg. It's so very flexible, so it's constantly developing. We've always been able to easily manipulate the database to do whatever we needed to do.

Of course, there are two sides to any database - the technology side and the content side. The content side is really what ultimately determines the value of any database. Bob Gex, then chief librarian at SLAC, was responsible for the content side, along with our long-time and remarkable preprint librarian, Rita Taylor. And from the beginning, there have been major contributions of content from our DESY partners.

FM: When did the Web first arrive at SLAC?

LA: On December 12, 1991, the first Web server in the U.S. was established at SLAC to help improve access to the SPIRES-HEP database. For a number of years, SLAC had been providing remote access to the SPIRES-HEP database. In fact, by the time we launched our first Web server, the SPIRES-HEP database had nearly 5,000 registered users in 40 countries.

George Crane had developed a remote interface, initially using BITNET. Researchers could send a search query to SPIRES-HEP via an interactive messaging tool, which was sort of like the instant messaging tools of today. Or they could request information via e-mail. When Internet access came along, the interactive messaging was no longer available; researchers had to use e-mail. Researchers were less satisfied with the e-mail interface; they liked the instant response system better.

On December 12, 1991, the first Web server in the U.S. was established at SLAC to help improve access to the SPIRES-HEP database.

FM: Was this a motivating issue when you first thought of setting up a Web server?

LA: Well, the real motivating issue was a project that we been working on earlier. In 1989-1990, I was helping set up a library at the Superconducting Supercollider (SSC) project in Texas. Pat Kreitz, the SCC librarian, wanted to use the SPIRES-HEP database and wanted to present it to the SSC community in a seamless way within a graphical user interface. SSC had funded the development of an X-Windows interface to do this, but as soon as I saw the Web in action, I knew that it was a fast, cheap solution.

I can still remember the day Paul Kunz appeared in my office after his trip to CERN. He showed me the Web and we started moving right away. I scurried around and got some accounts that were needed. Paul and Terry Hung set up the server on our mainframe. George Crane was able to easily extend remote SPIRES to talk to the new interface. And I was able to make the SPIRES-HEP database write HTML.

I've always felt deeply indebted to Paul for spotting the potential of the Web for our situation and for helping develop the first Web server at SLAC, which proved that potential.

FM: This sounds like it was really a team effort.

LA: It was! While I was working on this project - and for most of my other projects at SLAC - I was constantly seeking input and advice from colleagues. There was almost always someone around who understood the most arcane, technical issues and was willing to help or to give advice. Sometimes conflicting advice!

FM: Did this teamwork model work?

LA: Yes, it was the only way any of these projects could continue. None of us got raises or payment for this work. We did some of it on our own time. We knew that there was value in the work - and it was fun! (I hope my colleagues would agree.)

By February 1992, I had twisted enough arms to start an ad hoc Web development and support group, the WWWizards. The initial group was made up of Mark Barnett, George Crane, Tony Johnson, Joan Winters, Bebo White, all of whom brought important skills to the project.

The original WWW Wizards at SLAC visit Paul Kunz. Left to right: Louise Addis, George Crane, Tony Johnson, Joan Winters, Paul Kunz, and a NeXt computer (missing: Bebo White and Mark Barnett). Photograph courtesy of the Stanford Linear Accelerator Center, Archives & History Office, Stanford, Calif.

Again, I have to mention that in the beginning this was not part of our regular jobs. Luckily, the SLAC Computing Center was quite supportive. Les Cottrell, the assistant director of computing, could certainly have discouraged the computing staff from participating. Instead, they were able to work on the WWWizards team and to support this unsupported software. And, they helped run the server on their computer!

FM: Did you have any idea then of what the Web would become?

LA: No! I was just hoping it would survive at SLAC. For a long time, the Web suffered from the stigma of being unsupported software.

My goal was simply to provide better community access to the particle physics literature via SPIRES-HEP.

FM: Was SLAC a good environment for developing new tools, such as SPIRES-HEP or the Web?

LA: Yes! We always had the ability to respond quickly and be flexible. This was particularly instrumental in the development of the Web. When Paul Kunz came back from Europe and said, "Let's do this," the Library could move ahead.

And as soon as we were on the Web, we were able to start linking the records in our database to the TeX source at Los Alamos National Laboratory (LANL) where, in August 1991, Paul Ginsparg had started the first e-print server. Paul, by the way, almost single-handedly made another kind of revolution in the way scientific literature is handled ... but that's a whole other story.

As soon as we were on the Web, we were able to start linking the records in our database to the TeX source at Los Alamos National Laboratory.

FM: How were you able to just jump right into this new technology?

LA: First of all, our faculty were used to using unsupported software. We didn't have manuals or training classes, but we did have all kinds of skills and knowledge available for something interesting. The Web was done entirely with volunteer labor at first.

Generally speaking, the Lab encourages experimentation, sometimes officially and sometimes not. The Web server project was something that the SLAC Library could jump right into because of our track record with SPIRES-HEP and our environment. All that had gone before made it easier to get management approval. Plus, the chief librarian, Bob Gex, was terribly supportive. I was lucky in that regard.

MH: Do you feel that development of resources such as SPIRES-HEP has gone faster with the implementation of Web technology?

LA: Oh yes. In fact, this was one of the things that really sold me on the web - it was easy! The underlying Web server and Web browser programming is more technically complex, but creating the resources is much easier. For example, the x-windows project relied on programmers who could manipulate a more complex system. On the other hand, the Web allows for easy development of resources; you can make the page look exactly the way you want without deep programming.

FM: The particle physics world seems tailor-made for the World Wide Web - or should I say that the Web was tailor-made for the physics world?

LA: Absolutely! There is the issue of the need for rapid communication in a field that has a slow publication schedule. Additionally, the particle physics community is made up of very large groups of scientists all working on the same project, but from remote, diverse sites.

Also, these folks are used to trying something new. When we notified our registered users about how to get a free browser from Tim at CERN, many did right away. They wanted to get to our database, so they were really motivated to get that browser and get on the Web.

This helped spread use of the Web in the particle physics community and it also helped people learn about the archives at LANL.

FM: What about the democratizing of the publishing process?

LA: Gray literature is a very important part of the particle physics world. One of the great things about the Web is that it democratized access to these resources. Before electronic communication people who were further away from population centers or technology centers were really behind. They had no way to find out quickly about research.

Gray literature is a very important part of the particle physics world. One of the great things about the Web is that it democratized access to these resources.

Before machine-readable, full-text e-prints were available, paper preprints were mailed by the author or the institution only to major institutions. And researchers at smaller institutions didn't receive the preprints. The only way to find out about these articles was through a preprint list such as the one SLAC published weekly. The researcher then had to request the actual article through the mail. This was too slow.

FM: And the interactive nature of the Web?

LA: In large particle physics collaborations, the experiment may take years from conception through publication. And there may be hundreds of physicists around the world working on these projects. They are all producing materials that need input from others.

Many papers are almost dialogues. Papers go through revisions. Discussion is a big part of the process. Physicists really needed that ability to interact quickly.

The Web was a revolution!

FM: Back to your career development, did you feel more affiliated with the library community or the particle physics community?

LA: Much more affiliated with the particle physics community. Our issues were so different from what others in most libraries were experiencing. In a special library, you have to know your community and listen to them. They're the people that need the library.

FM: Doesn't this apply to all libraries?

LA: More so to special libraries, which are vulnerable and can be more easily closed.

FM: What do you think is the response to this vulnerability?

LA: We need to apply other skills and expertise in order to allow the library to succeed. You almost never have all the skills you need. You have to find them elsewhere. For example, at the SLAC Library we wanted to make PostScript files and graphics available via the Web, rather than just linking to the TeX sources. So, when Tony Johnson developed Midas, the first GUI browser that was bug-free enough for us to use, he added capability for reading and displaying postscript. This was the kind of development done on someone's personal initiative. This is skill and interest that you cannot buy. Also, Tony had the breadth of knowledge of a particle physicist's needs and a willingness to tinker.

FM: So, we're back to an environment of experimentation ...

LA: Yes, a lot of environments don't make it easy to go outside or take risks and that's what is required to be visionary. You can't pay attention to things like job descriptions.

FM: Is there anything that concerns you about the development of the Web?

LA: I'm concerned about commercialization. The Web has been a democratizing place. Now, we're moving toward large entities controlling most content - or the content that most people see. You can get to other sites, but these services shape your experience. AOL tells you the story of the day and offers you its preferred links.

On the one hand, anyone with a little curiosity can find things they'd never dream of having access to. But portals do control what a lot of us see - or have time to see! But this has always been true of the media; the Web is just another instance of this.

FM: What about the issue of the digital divide? Or does this problem not exist in the academic community?

LA: Actually, this has been a continuing problem for some of the remote or less prosperous regions of the world. In the U.S., a small physics department in a small school may have a slow line or slow equipment.

When I was at SLAC, we always created pages in formats that were much simpler than what was technically possible. Harv (Hrovje) Galic, the HEP database manager at that time, originated much of our early Web interface. Harv was adamant about trying to have pages that could be easily read in some of the less prosperous parts of the particle physics community. Harv was from Zagreb and had a real awareness of the constraints faced by many of his colleagues.

In fact, we were occasionally ridiculed because our pages didn't have graphics; but this was done purposely. The main reason was because many of our users didn't have the capability to see graphics - or graphics were an impediment on a slow system.

The digital divide is real. It's less of a problem in academic communities where the issue is fast or slow connection. Or good network support and stability. In society as a whole, the issue is having any access at all.

FM: Finally, do you think many folks will be surprised to learn that the first Web server in the U.S. was installed to help support a library project?

LA: Not at all! To me, it's really significant that our own early success story on the Web was driven by the need to bring a large body of carefully organized bibliographic records, SPIRES-HEP, to its worldwide audience!

The Web has certainly proven to be a shiny but challenging new tool for librarians to use in their crucial role as collectors, organizers, preservers, and presenters of information. Librarians will continue to play a critical role in helping to organize and provide access to information. Perhaps an even larger and more important role now that we're in the "wild, wild west" era of Web content development. FM

Notes

1. Tim Berners-Lee, 1999. *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by its Inventor.* New York: Harper SanFranciso, p. 45.

2. Berners-Lee, p. 46.

Editorial history

Paper received 13 April 2000; accepted for publication 17 April 2000.

Contents Index

FM Interviews: Louise Addis by Melissa Henderson
First Monday, volume 5, number 5 (May 2000),
URL: http://firstmonday.org/issues/issue5_5/addis/index.html

DOT-COM

Archivists try to save historic Web pages

BY THE ASSOCIATED PRESS

Seminal Web pages are vanishing before they can be saved, a situation a small group of local scientists and students is trying to remedy.

Many experts see early Web pages as similar to some of our nation's most important documents. Unlike those documents, which are safely on paper, a Web page exists only in the digital world. If they aren't archived, they can disappear before your eyes.

"They don't give any warning they're going to corrupt," said archivist Jean Deken, who works at the Stanford Linear Accelerator Center. "They simply very quietly pass out of usefulness."

Deken got involved a few years ago when she tried to find the nation's first Web page, which was created in an office at SLAC. Although no record of it existed, it was located on an old backup tape.

At the National Archives and Records Administration, they have been able to process only 100,000 electronic records over the past 25 years. But the Treasury Department alone generates nearly one million e-mail messages a year that experts think should be saved.

ATMs will play movie ads

BY THE ASSOCIATED PRESS

Movie previews are coming to an automated teller machine near you.

Providing a peek at Hollywood's latest blockbuster is just one of the new high-tech features that Wells Fargo Bank plans to unveil today as it rolls out a new breed of ATM.

The new machines represent the biggest step yet to transform ATMs from mere cash dispensers to full-blown multimedia centers.

The new machines are designed to duplicate the look of a personal computer screen by offering appealing graphics and icons that open new windows with a touch of the finger.

Operating the new ATMs doesn't require new skills, even though Wells is using powerful new technology, including Intel Corp.'s Pentium III microprocessors, to upgrade its ATMs.

The ATMs aren't connected to the Internet, but Wells calls them "Web-enabled" because the technology is now in place that could eventually let customers buy products at the terminals.

Animal Coins of 2000

(cont'd. from page 6)
to manage production of rf components built at SLAC while LLNL's Mark Franks continued to oversee the Livermore fabrication and industry contracts for the e-beam welding.

SLAC's Jim Judkins and Bob Rimmer from Berkeley performed the final frequency tuning at Livermore and in July this year the first production cavity was flow-balanced by LLNL's Ed Thomas and delivered to SLAC. Mark Hoyt and Jeff Jones fitted window and coupler components and HOM loads provided by engineers Karen Fant and Mike Neubauer and designer Tim Montagne. A new tuner was fitted courtesy of Heinz Schwarz and designer Ed Evans. The cavity was baked in an oven provided by Mike Zurawel and Jay Venti of MFD and installed in the test bunker by Jim Judkins, Alan Hill, Alden Owens and Marcus Gorgees. The cavity processed quickly up to full power using a computer control system put together by Alan Hill, and on September 26 satisfied the criteria of a major project milestone, tracked by the DOE, for a "full rf cavity system integration test," once again showing that the project is right on track.

The development of the cavity from its earliest inception to the current production status would not have been possible without the help of many more people than it is possible to mention here by name, but they can all be proud of their contributions when the machine comes on line in 1998.

SLAC First US WWW Site, Not Fermilab

IN THE AUGUST 16 issue of *FermiNews*, the front-page headline reads "Birthplace of the Web" which one could interpret as meaning that the first US Web site was at Fermilab. Not so, according to SLAC experts Les Cottrell, Paul Kunz, and Joan Winters.

SLAC was the first Web site in the US, with a verified date of December 12, 1991, according to Robert Calliou of CERN. Fermilab joined in June 1992.

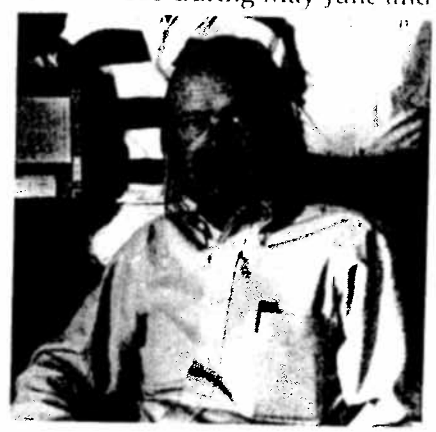
Close reading of *FermiNews* shows that the newsletter is crediting the field of high energy physics as the impetus for the spread of the Web to the US and on that point we can agree.

Around the Globe with C++

C++ IS AN OBJECT ORIENTED programming language, first released in 1985. At that time, FORTRAN was the programming package used by the scientific community. In 1989, Paul Kunz (Group B) and Bill Atwood (Group EK) were among the few SLAC staff who started learning object oriented programming to solve a lot of difficulties in maintaining large code. FORTRAN is still heavily used by physicists although object oriented programming has a large following in the industrial community.

Then BaBar submitted their *Technical Design Report* in March 1995 and announced that their software would be based on object oriented programming. Out of over 500 collaborators, less than a dozen knew how to use C++. So, how do you teach technical information to almost 500 people worldwide? Do they learn it slowly by actually using it, or take a course? Hands-on training was too difficult to oversee or to pick up the needed skills. So, then how to train—textbook or workshops by professionals?

It was decided that the best way was to have a course taught by a physicist for physicists to get the right flavor. Then examples would be directly from problems that they were used to dealing with. Since Kunz had been working with C++ for the past 7 years, the BaBar Collaboration invited him to design and teach a course, both to teach C++ and to show that it was the right programming decision to make. Not surprisingly, there was reluctance to move to C++ due to familiarity with FORTRAN. The first class was held at SLAC in 1995 during May-June and met once a week for six weeks. It took Kunz the rest of each week to prepare the lectures. It can be noted here that the lecture notes have not changed significantly since this preparation was done. He quickly followed up with lectures at Lawrence Berkeley National Laboratory and Lawrence Livermore National Laboratory, then back for two more SLAC lectures. From May through September, 150 physicists attended his classes.


Paul Kunz

BaBar collaborators are not the only ones interested in learning C++ in this interactive style; over the past year, laboratories worldwide have invited Kunz to their locations. To date, over 750 physicists from six countries have attended his week-long lecture. Kunz, itinerant physicist, is now a frequent flying teacher.

ACKNOWLEDGMENTS

SAS Programming: Lois White

CANDO Database: Teresa Downey

TRICKLET SNMP routines: Don Pelton

The members of our Network Development and Network Operations Groups whose comments from the battlefield have been invaluable.

Thank you for your time.

Overheads:

ftp.slac.stanford.edu

get ~ftp/pub/cal/SNMP-9-23.ps

DISCLAIMER

" The opinions expressed here are my own, and not those of the Stanford Linear Accelerator Center, Stanford University, the Department of Energy, or the United States government"

Connie Logg

Thoughts for the Future

What we have is working very well for us now--152 interfaces.

- 3MB per day data
- 300MB for graphs

However there are scalability problems for some technologies - for example:

- Multiport repeaters with 9 ports...90 of them. That is 810 interfaces/ports...
- 10baset units with 13 interfaces/ports...20 of them. That is 260 interfaces/ports...for just 20 units!

This requires the development of a different data collection and data display scheme. I am working on that now.

- SNMP traps (unreliable)
- collect data less frequently
- monitor, but collect data only when problems are detected
- Don't generate all graphs & don't save all data

Directories & disk management

What Vendors Can Do to Help

- IMPLEMENT STANDARD MIBS - document where to get the source
- Implement the full port statistics
- Provide simple explanations of the terminology - do not assume that the customer is an expert.
- Define explicitly what parts of the MIBs are implemented and not implemented.
- Implement error counters - DO NOT SKIMP HERE
- Provide step by step configuration information and a GUI to do it if possible.

You, as a manufacturer/vendor only have to do this once. IT IS COST EFFECTIVE and by doing it thoroughly you will save yourselves a thousand times (or more) the cost of doing it right, because:

- Customers will need less support...they will be much more self-sufficient
- They will feel good about themselves and your products and tell their friends how easy your products are to understand and use.

Traceroute and Ping for Response Time Analysis

Perform a traceroute to the complaining node

Analyze the response time and packet loss:

NODE	#pings	#loss	%loss	min	max	avg	#loss	%loss	min	max	avg
192.68.191.2	525	0	0.0%	2	7	2	0	0.0%	5	11	6
134.55.4.226	525	0	0.0%	6	18	9	0	0.0%	21	34	24
134.55.12.98	525	0	0.0%	7	19	9	0	0.0%	23	34	25
134.55.12.65	525	0	0.0%	8	17	10	0	0.0%	26	37	29
134.55.12.225	525	0	0.0%	10	24	13	1	0.2%	34	50	38
192.188.35.1	525	30	5.7%	11	24	15	34	6.5%	37	48	40
128.115.237.254	525	46	8.8%	13	24	17	42	8.0%	43	59	48
128.115.245.2	525	41	7.8%	13	44	21	43	8.2%	45	64	50
GEM1.LLNL.GOV	525	43	8.2%	14	25	18	35	6.7%	52	66	56
			100 byte packets					1000 byte packets			

From the above table one can pinpoint where the loss and slowdown starts.

Summary and Conclusion

For the maintenance of large dynamic networks, extensive monitoring of the network equipment and media is necessary to provide for proactive management and maintenance

Monitoring a large network entails looking at a very large amount of data. This process must be automated to make it feasible.

To handle the volume of data involved, it must be reduced so that the data indicating potential problems can be efficiently and effectively examined daily or as frequently as desired.

Presentation via the World Wide Web, as described here has been very effective in providing the information necessary to maintain the quality of the SLAC networks.

All the data collected is readily available at our fingertips, if it is needed. However, our daily or more frequent (if necessary) monitoring is reduced to only looking at some short reports that point out potentially troubled areas that may exist.

Hints for Troubleshooting Network Performance Problems

Explore your data - Feed it into a spreadsheet and plot it

Look for correlations - Identify changes in baselines and compare to other changes in network activity and the environment:

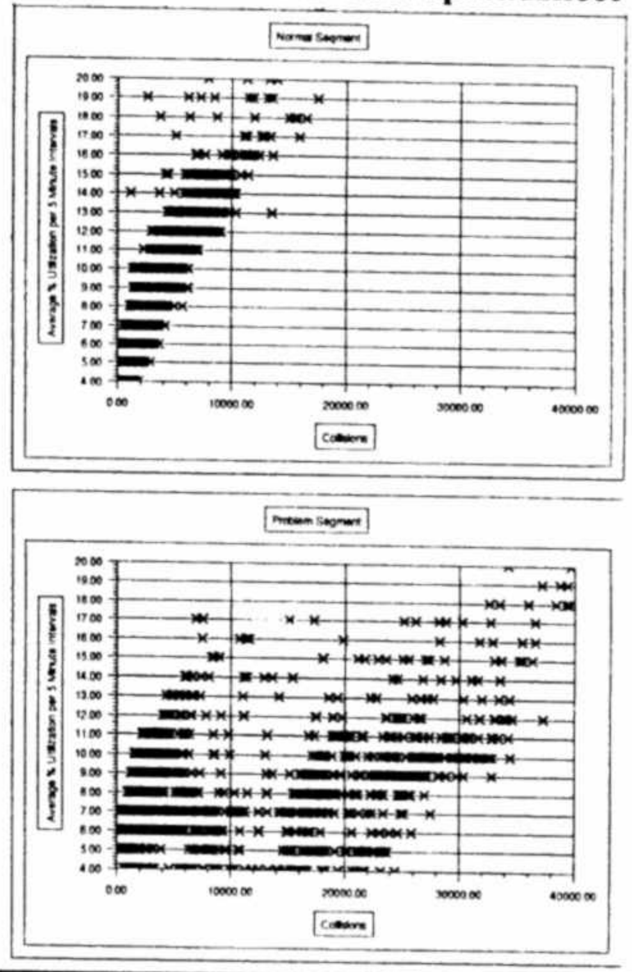
- variations in the percentages of protocols on a net
- changes in utilization by specific nodes
- weather changes (temperature, rain)
- changes in packet size distribution
- changes in baselines after topology changes
- application and system software changes

For complaints of slow network response - Examine response time of the various nodes (e.g. routers, bridges, etc.) along the way

Examine the traffic patterns - Are the nodes located on the correct networks (see page 22, Top10 Talkers)

USE YOUR IMAGINATION!

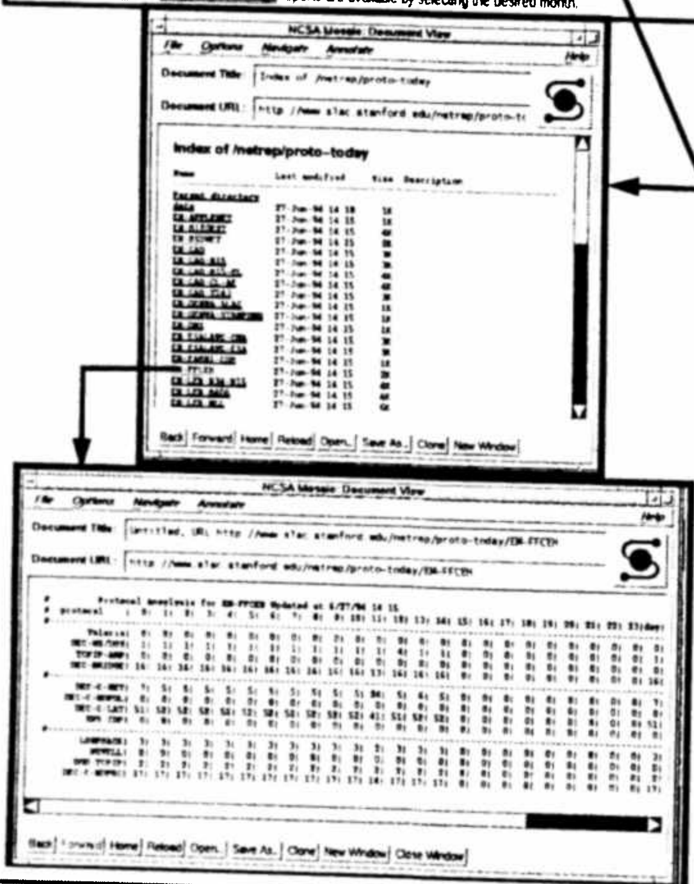
Explore Your Data via a Spreadsheet



Protocol Analyses

PROTOCOL ANALYSES -

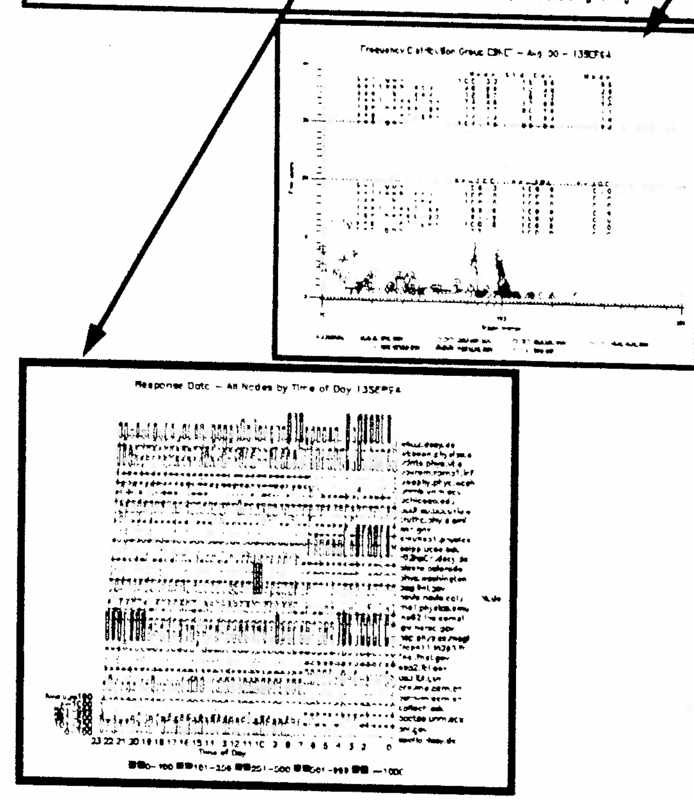
- Daily summaries are available for today and yesterday, as are the individual network reports for today and yesterday.
- The raw data and previous percentage reports are available by selecting the desired month.



Ping Analysis and Reports

PINGING - Various nodes on the network are pinged periodically to sample for packet loss and response time

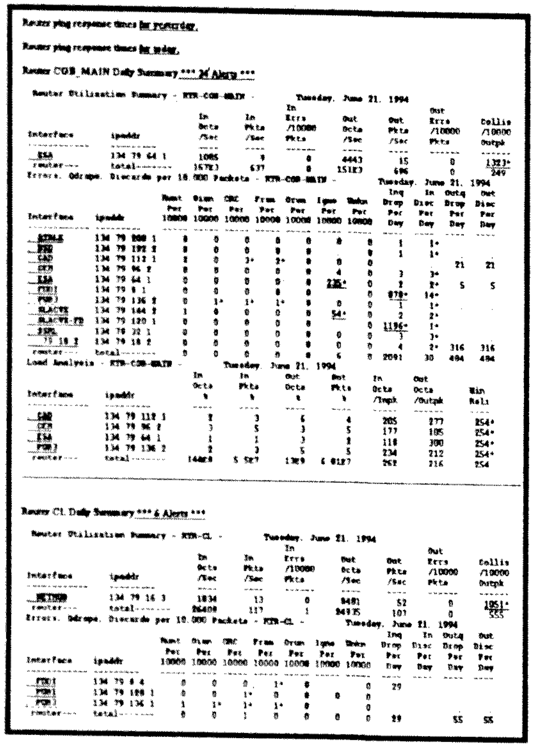
- Pinging of collaborators' offsite nodes samples their response time and availability. Plots of additional statistics and individual node responses are also available. Particularly enlightening plots are the frequency distribution of response times for nodes on ESNET, INTERNATIONAL, and NSFNET.
- Pinging of critical servers samples their response time and availability. Plots of additional statistics and individual node responses are also available. The frequency distribution of response times for critical server nodes is particularly enlightening.



Quick Review - Router Alerts Summary

For a *QUICK REVIEW* of yesterday's happenings and data analysis and reduction see:

- The Summary of Potential Problems for yesterday. This report contains exceptions noted in the Ethermeter and Bridge reports plus critical node outages. Hot links are available to reports and plots providing further information.
- The Router Alerts Summary for yesterday, gives a tabular report of router statistics that are outside thresholds.
- It is often a good idea to review the last 40 PROBTRAK entries to see what problems have been reported recently.
- The Change Log also provides a log of incidents and changes.

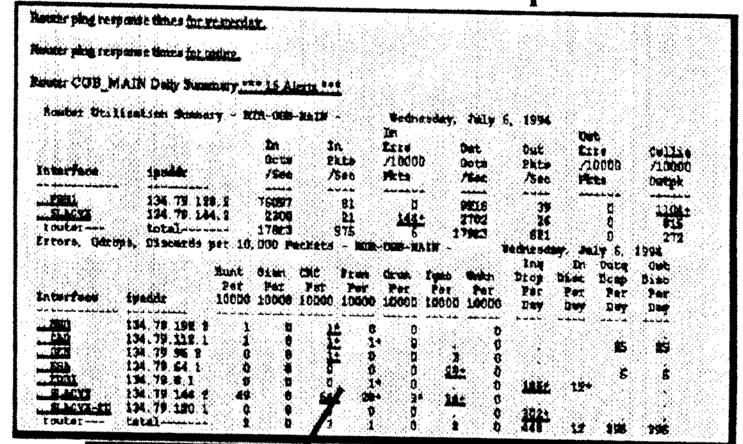

Router ping response times for yesterday.

Router ping response times for today.

Router CGB_MAIN Daily Summary *** 24 Alerts ***

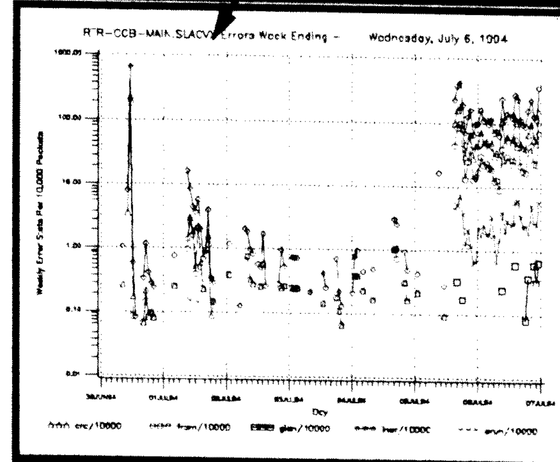
Router CL Daily Summary *** 6 Alerts ***

Selected Router Graph


Router ping response times for yesterday.

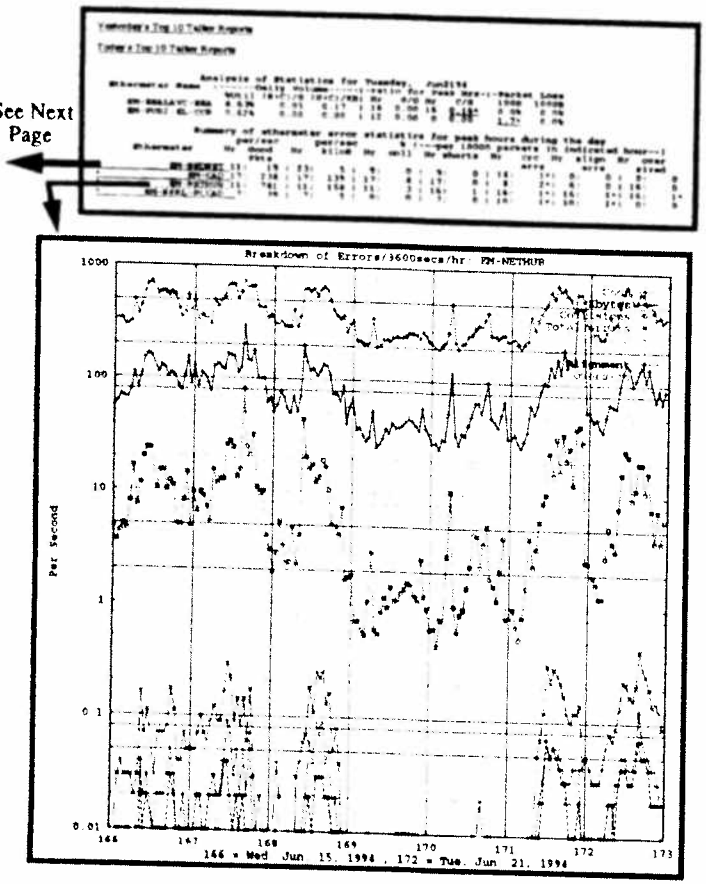
Router ping response times for today.

Router CGB_MAIN Daily Summary *** 15 Alerts ***



Selected "EM-NETHUB" Data Graph

See Next Page



Subnet Menu

- Quick Guide.
- View network maps.
- List of subnets.

1 ROUTER found. 1 ETHERMETER found. 0 LAN-BRIDGEs found.

Contents

Today's Reports

Today's Ethermeter Volume graphs
Today's Ethermeter Error graphs
Today's Ethermeter Protocol data
Today's Ethermeter Top 10 data
Today's Router Volume graphs for physical router= ROUTER2
Today's Router Error graphs for physical router= ROUTER2
Today's Router Miscellaneous graphs for physical router= ROUTER2

Yesterdays Reports

Yesterday's Ethermeter Volume graphs
Yesterdays' Ethermeter Error graphs
Yesterday's Ethermeter Protocol data
Yesterday's Ethermeter Top 10 data
Yesterday's Ethermeter Packet Loss data
Yesterday's Router Volume graphs for physical router= ROUTER2
Yesterday's Router Error graphs for physical router= ROUTER2
Yesterday's Router Miscellaneous graphs for physical router= ROUTER2

Last 14 Days Reports

Last 14 days' Ethermeter Volume graphs
Last 14 days' Ethermeter Error graphs
Last 14 days' Ethermeter Peak graphs
Last 14 days' Router Volume graphs for physical router= ROUTER2
Last 14 days' Router Errors graphs for physical router= ROUTER2
Last 14 days' Router Miscellaneous graphs for physical router= ROUTER2

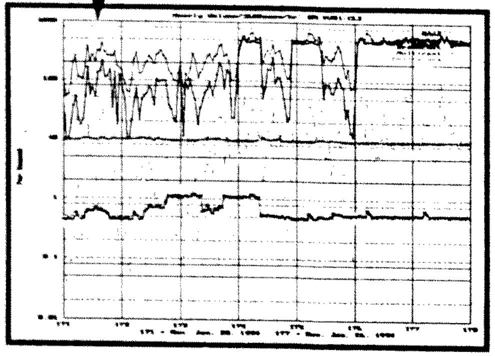
Last 28 Days Reports

Quick Review - Continued

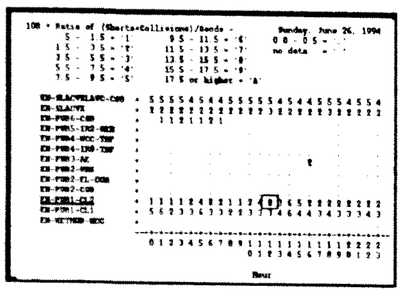
Summary of Potential Problems

Top10 talkers for yesterday

```
              Analysis of Statistics for Sunday,  Jun2694
Ethermeter Name  |------Daily Volume-----|-ratio for Peak Hrs-|-Packet Loss
                 %Util  (S+C)/G  (S+C)/KB|  Hr   S/G  Hr   C/G    100B   1000B
     EM-PUB1-CL2  38 94*   0.04    0.04  |   6  0.00  12  0.15*   0.0%    0.0%
EM-SLDLAVP-CL-IR10 4 76%   0.03    0.11  |  12  0.00   9  0.07    0.0%    2.9*
        EM-TEST1  12 81*   0.03    0.07  |  10  0.00   9  0.04    4.6*    2.1*
        EM-TEST2  12 81*   0.03    0.07  |  10  0.00   9  0.04    2.1*    0.0%
```



see next slide
Top10 Talker Table



Quick Review - Continued
Top 10 Talker Table

Top 10 Talkers per hour for 94/06/26 00:00:00 to 94/06/26 23:59:59 for /nfs/juno/u20/netdata/top10/data/em-pub1-cl2 jun94

Numbers represent the percent of the total packets a pair represents for that hour
Note that total represented by the top10 may not match the sum of the individual entries, due to roundoff

From node	To node	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
EANEXTN	ROUTER6						1																		
EBNEXTW	ROUTER6						4																		
EBNEXTK	KAON										0														
EBNEXTK	IGUANA														0										
EBNEXTK	ROUTER6					2	2																		
ZONEXT1	ROUTER6					4																			
CARDINAL	ROUTER6					1																			
THNEXT	ROUTER6						5																		
LIBNEXT	ROUTER6							0																	
KPLUS	GEMINI	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SNUFFY	ROUTER6								0	1															
XSLACLIB1	ROUTER6														0										
KERMIT	ROUTER6	0	0	0	0			0	0	0	0	0				0				0			0	0	0
GEMINI	KPLUS	0	0	0	0			0	0	0	0	0				0								0	0
GEMINI	ROUTER6							1																	
RAVEL	ROUTER6												10	19	2	14	5								
HERON	ROUTER6						1											1							0
DALI	ROUTER2																		1		1				
AVOCET	ROUTER6	0																							
EGRET	ROUTER2											1					1	1	1	1	1	2			
ALBATROSS	ROUTER2																1	1							
AIXOL01	ROUTER6																			1					
WGYMAC2	ROUTER6		0	0				0	0	0	1	0	1	1	0	0				1	1	1	1	1	
PIXEL	ROUTER2	1			3								1											0	
TURKEY	ROUTER6		1	0														1	1	1					
AIXPUB02	ROUTER6	0	0	0	0			0	0	0	0	0													0
HAWK	ROUTER6	26	26	26	24	22	17	23	24	25	26	25	22	19	23	19	23	23	23	24	24	24	24	24	29
SLACSB	ROUTER6																						0		
ROUTER6	AIXPUB02	1	1	1	1	1		1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1
ROUTER6	HAWK	62	62	63	59	52	40	55	59	62	62	61	52	46	55	45	55	57	56	59	59	58	58	57	60
ROUTER6	09000700001d		0	0	0				0	0	0											0	0		
ROUTER6	SLACSB																		1						
ROUTER2	EBNEXTW						4																		
ROUTER2	EBNEXTK					2	2																		
ROUTER2	ZONEXT1					4																			
ROUTER2	THNEXT						5																		
ROUTER2	XSLACLIB1														1	0									
ROUTER2	RAVEL												2	3	6	12	1								
ROUTER2	GOYA												1	0											
ROUTER2	DALI													1					1		1				
ROUTER2	EGRET											1					1	2	1	1	1	2			
ROUTER2	ALBATROSS					1											1	1							
ROUTER2	PICASSO																					1	1	3	2
ROUTER2	PIXEL	1			3								1	0										1	
ROUTER2	09000700001d																								0
ROUTER2	SLACSB																				0	1	1		
Top 10 Total %		92	93	92	92	88	79	81	87	91	91	91	90	89	89	91	88	87	86	90	88	89	87	88	89
From node	To node	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

HOUR

The WWW Network Reports Page

Quick Guide to SLAC's Networking Reports and Data – Overview

There are various [illegible] applied to the daily data analysis and reduction for the extraction of "alertable" situations.

For a *QUICK REVIEW* of yesterday's happenings and data analysis and reduction see:

- The [illegible] for yesterday. This report contains exceptions noted in the Ethermeter and Bridge reports plus critical node outages. Hot links are available to reports and plots providing further information.
- The [illegible] for yesterday, gives a tabular report of router statistics that are outside thresholds.
- It is often a good idea to review the [illegible] entries to see what problems have been reported recently.
- The [illegible] also provides a log of incidents and changes.
- One can also review the data for segments of the network by [illegible].

The netnews group [illegible] gives a list of scheduled outages.

The [illegible] are from the most recent H.O.T. meeting.

The network [illegible] provide skeleton maps of the network layout.

[illegible] are the primary instrument for measuring ethernet traffic and errors on networks at SLAC. The following are the tabular reports and graphs which are updated hourly for today's data and daily for the rest.

- Tabular report of ethermeter statistics (non-errors) for [illegible].
- Tabular report of ethermeter errors for [illegible].
- Ethermeter Volume graphs for [illegible] or [illegible] or the [illegible] or the [illegible] or the [illegible] or the [illegible].
- Ethermeter Error graphs for [illegible] or [illegible] or the [illegible] or the [illegible] or the [illegible] or the [illegible].
- Ethermeter peak graphs for [illegible] or [illegible] or the [illegible] or the [illegible] or the [illegible] or the [illegible].
- The ratio of the number of shorts+collisions to good-packets [illegible] provides a quick view of network hot spots for yesterday.
- The TOP-10 talker reports show the hourly top 10 talkers on each net in a tabular form for [illegible] and yesterday.
- All ethermeters are pinged periodically to sample for packet loss and response time. [illegible] reports are available for yesterday. Response time reports are available for [illegible] and for the [illegible].

[illegible] We collect data hourly from the NAT bridges interfaces in the network.

- Tabular report of the bridge statistics for [illegible].
- NAT bridge data graphs are available for [illegible] or [illegible] or the [illegible] or the [illegible] or the [illegible].

[illegible] We collect data hourly from the CISCO router interfaces in the network.

- Cisco router interface graphs are available for [illegible] or [illegible] or [illegible] or [illegible] or [illegible] or [illegible].
- All the Cisco router interfaces are pinged periodically to sample for packet loss and response time. Reports are available for [illegible] (updated hourly) and [illegible].
- The [illegible] from the router data collection is also available.

FDDI RING ANALYSES – Under Development

- Hourly updated summaries for [illegible] are available for the main SLAC ring.
- Daily summaries are available for [illegible] for the main SLAC ring.

[illegible]

- Daily summaries are available for [illegible] and [illegible], as are the individual network reports for [illegible] and [illegible].
- The [illegible] and [illegible] reports are available by selecting the desired month.

[illegible] – Various nodes on the network are pinged periodically to sample for packet loss and response time.

- Pinging of collaborators' [illegible] samples their [illegible]. [illegible] of additional statistics and individual node responses are also available. Particularly enlightening plots are the frequency distribution of response times for nodes on [illegible] and [illegible].

Quick Review - Summary of Potential Problems

For a *QUICK REVIEW* of yesterday's happenings and data analysis and reduction see:

- The Summary of Potential Problems for yesterday. This report contains exceptions noted in the Ethermeter and Bridge reports plus critical node outages. Hot links are available to reports and plots providing further information.
- The Router Alerts Summary for yesterday, gives a tabular report of router statistics that are outside thresholds.
- It is often a good idea to review the last 40 PROBTRAK entries to see what problems have been reported recently.
- The Change Log also provides a log of incidents and changes.

Yesterday's Top 10 Talker Reports

Today's Top 10 Talker Reports

```
                     Analysis of Statistics for Tuesday,  Jun2194
  Ethermeter Name  |------Daily Volume-----|-ratio for Peak Hrs-|-Packet Loss
                   %Util (S+C)/G (S+C)/KB| Hr   S/G Hr    C/G    100B  1000B
    EM-ESALAVC-ESA  4.63%   0.05    0.17 | 16  0.00 16  0.16*   0.0%   0.0%
    EM-PUB2-EL-CCR  0.62%   0.00    0.00 | 12  0.00  0  0.00    1.7*   0.0%

             Summary of ethermeter error statistics for peak hours during the day
                  per/sec       per/sec       % (---per 10000 packets in indicated hour--)
     Ethermeter   Hr  Good   Hr   kiloB  Hr  coll  Hr shorts  Hr   crc  Hr  align  Hr  over
                      Pkts                                         errs       errs       sized
*      EM-BSDNET 11    19 | 23     5 |  9    0 |  9     0 | 14    1*|  0     0 |  0    0
*         EM-CAD 17   234 | 17   139 | 17    4 | 17     0 |  8    2*|  6     0 | 16    0
*      EM-NETHUB 11   741 | 11   154 | 11    3 | 16     1 | 16    1*| 16    1*| 16    1*
*   EM-SSRL-PCCAD  7    39 |  7     5 |  0    0 |  7     0 | 10    1*| 10    1*|  0    0

             Summary of Statistics for peak hours during the day
                  per/sec (---------per 10000 packets in indicated hour----------)  per/sec
     Bridge & if   Hr  Good   Hr buffer Hr  crc   Hr  align  Hr  ctrl  Hr  coll  Hr  multi
                       Pkts      ovflow     errs       errs       ovflow               casts
* EB-CAD-COB-B209-1 17  223 |  0    0 |  8   2*|  6    0 |  0    0 |  0    0 | 11    6
*    EB-CAD-CL-AE-1 17  198 |  0    0 |  8   2*|  6    0 |  0    0 |  0    0 | 11    6
*     EB-CAD-EL1-2 16   155 |  0    0 |  8   4*|  6    1*|  0    0 |  0    0 | 12    6
*  EB-NETHUB-COB-1 18   515 | 22   20*| 18    0 | 17    0 |  0    0 | 23    0 |  8    3
* EB-NETHUB-IR10-1 16   689 | 22    2*| 18    0 | 19    0 |  0    0 | 23    0 |  8   13
*    EB-SLC-MCC1-1 23   783 | 22    3*|  0    0 |  0    0 |  0    0 |  0    0 | 13   24
*  EB-SSRL-PCCAD-1  7    13 |  0    0 | 11   4*| 11    3*|  0    0 |  0    0 |  9    0
* EB-SSRL-PCCAD2-1 10     2 |  0    0 |  9   4*|  0    0 |  0    0 |  0    0 |  9    0

                            Node Outage Summary

QMAIL  134.79.209.10 outage:
   from TUE_JUN_21_13:51:12_1994 to TUE_JUN_21_13:54:14_1994,    3.03 min.
QMAIL  134.79.209.10 outage
   from TUE_JUN_21_23:03:00_1994 to TUE_JUN_21_23:04:01_1994,    1.02 min
QMAIL  134.79.209.10 outage
   from TUE_JUN_21_23:11:04_1994 to WED_JUN_22_02:02:03_1994,   48.93 min
SLACVX  134.79.120.12 outage:
   from TUE_JUN_21_14:19:26_1994 to TUE_JUN_21_14:20:24_1994,    0.97 min.
SLACVX  134.79.120.12 outage.
   from TUE_JUN_21_14:21:30_1994 to TUE_JUN_21_14:22:32_1994,    1.03 min
SLACVX  134.79.144.12 outage
   from TUE_JUN_21_14:19:26_1994 to TUE_JUN_21_14:20:24_1994,    0.97 min.
SLACVX  134.79.144.12 outage.
   from TUE_JUN_21_14:21:29_1994 to TUE_JUN_21_14:22:32_1994,    1.05 min.
SERV03  134.79.16.9 outage
   from TUE_JUN_21_08:16:16_1994 to TUE_JUN_21_08:34:03_1994,   17.78 min.
SERV03  134.79.16.9 outage
   from TUE_JUN_21_14:25:16_1994 to TUE_JUN_21_14:47:27_1994,   22.18 min.
```

Thresholds - continued

When the router data is analyzed and reduced, values which exceed these thresholds are alerted:

- total interface input errors over 1 in 10000 packets
- collision rates over 1000 in 10000 packets
- crc errors and alignment errors > 1 in 10000 packets
- buffer overflows and controller overflows
- in q and out q drops and discards
- ignored packets
- router interface ping packet losses > 1%

Futures:

- Alert on heavy users (maybe over 20% for the day or hour)
- Alert on unauthorized protocols

The Data Presentation

World Wide Web

We make heavy use of the World Wide Web in providing access to the hundreds of raw data files, graphs, and tabular reports.

The WWW default display is a listing of directory entries from which the user chooses the file to be looked at.

We have written our own display page (in a language called html) for providing access to the network data, graphs, and tables. It is entitled:

Quick Guide to SLAC's Networking Reports and Data

The Data Reduction Process

The data reduction process examines all the ascii reports which were generated by the analysis, and extracts the data which exceeded certain alertable thresholds.

The alerted data are then processed by programs which prepare them for presentation via World Wide Web.

This includes inserting into the extracted alerted data hypertext links to various tables and graphs which may provide more information about the alerts.

This ties together the few interesting graphs and tables which may be helpful in understanding an alert to the alert itself.

Network Data Thresholds

When the ethermeter data is analyzed and reduced, values which exceed these thresholds are alerted:

- crc errors and alignment errors > 1 in 10000 packets
- total utilization on a network over 10% for the day
- broadcast rate > 300 per second
- (shorts+collisions)/good_packets > .10
- packet losses from ping tests greater than 1% in a day

When the bridge data is analyzed and reduced,values which exceed these thresholds are alerted:

- crc errors and alignment errors > 1 in 10000 packets
- buffer overflows and controller overflows

The Data Analysis Process - Continued

Once a day (in the early am):

- The previous day's data is analyzed and summarized into several ascii files, with data points which exceed certain thresholds highlighted
- Graphs of data are generated for the previous day, the last 7 days, the last 14 days, and the last 28 days.

The data analysis process generates literally over thousands of data graphs and numerous ascii reports, most of which are uninteresting.

The Data Analysis Process

Ongoing during the day consists of:

- Generating hourly graphs and other displays of data collected so far "today"
  - Bridge interface stats,
  - ethermeter interface stats,
  - top10 talkers,
  - router interface responses,
  - subnet protocol analyses
- Generating alerts as potential problems are noticed
  - Duplicate IP addresses
  - Loss of connectivity between the NMS and a node
  - Data values exceeding thresholds (future)

Data Collection Control and Configuration

The data collection is driven by tables of equipment IP names extracted via SQL from CANDO.

Note that a database system is not necessary. These could be manually maintained tables.

The data collection is controlled by crontab entries

- most data collection takes place once an hour
- some is done more frequently

All data is written to ascii readable flat files.

Data Collection Via Ping

Ping:

- critical servers
- router interfaces,
- ethermeters
- off-site collaborators' nodes
- other interfaces/nodes as needed

Provides sampling information on:

- response times
- packet loss
- connectivity

The SNMP Data Collection - Cont.

- Types of data:
  - Interface stats
  - # good packets
  - # errors (# of types of errors)
  - # kilobytes
  - # pkts dropped, discarded, internal buffer overflows
  - one minute peak rates for the above
  - distribution of packet sizes
  - Top10 talkers
  - Protocol distribution on the subnets
  - ARP caches - look for duplicate IP addresses, unregistered IP addresses, ethernet/IP pair address changes (to update CANDO)

Raw Data Disk Resources

Ethermeters: 60 currently

ifs: 60 * 7300 = 438,000 bytes/per day

protocol data: 60 * 10000 = 600,000 bytes/day

top10 talker: 60 * 25000 = 1,500,000 bytes/day

Bridges: 27 * 2 = 54 interfaces

ifs: 54 * 2600 = 140,400 bytes/day

Routers: 38 interfaces

ifs: 38 * 8800 = 334,400 bytes per day

About 3 MB per day for raw data

Unix Based Data Collection and Analysis Tools

- IBM RS6000/370T and Sun SPARCserver 10/51
- **SNMP utilities (CMU, Tricklet, and vendors)
- **Ping
- **World Wide Web
- *Crontabs
- *Perl
- *SAS
- *Rexx
- *Gnuplot
- Traceroute
- NMS (Netview/6000 & DEC MSU)
- Framemaker
- Wingz
- Oracle and SQL
- CANDO (Computer and Networking Database in Oracle) - contains info on networked equipment

The Data Collection Process

Via SNMP:

- Currently collect data from:
  - Bridges
  - Routers
  - Ethermeters (NAT)
- Future (may require different collection/analysis scheme):
  - Hubs
  - Switches
  - Multiport Repeaters
  - Gigaswitch

SLAC Network Characterization Sept-94

Metric	Value
Number of on-site users	1800
Total Number of IP Nodes	1560
Macs & Other PC nodes	800
Unix Nodes	350
VMS Nodes	240
Xterminals	200
FDDI rings	2 (Will be 6 by end of 1994)
Routers / Bridges	7 (38 ifaces) / 27 (54 ifaces)
AppleTalk IP Gateways	26
Ethermeters	60
Segments / Subnets	480 / 24
Protocols	IP, DECnet/LAVC etc,Apple-talk, XNS, IPX
Operating Systems	Unixes, VMS, DOS, OS/2, Windows, MacOS, VM, Amiga, OSF/1...
Off-site Links	BARRnet, ESnet, Canada

Goals of Monitoring

- Measure and monitor the health of the network
- Become aware of problems which are developing and fix them before the users notice them
- Monitor the growth of network traffic and plan ahead for needed expansions
- Monitor the traffic patterns on the network to provide for optimal network design

Map of the SLAC Campus

2 mile long linear accelerator

Aerial View of SLAC
(adapted from photo by Joe Faust, 27 Feb 1987)

Research Areas

Nature of the SLAC Network

The SLAC network:

- is spread physically all over the SLAC campus
- dynamic - continually growing - by a factor of 2 in the last year
- has a frequently changing topology as proactive analysis shows hot spots
- 30 - 50 new IP addresses a month
- New technologies continually incorporated

Network Performance Monitoring and Analysis at SLAC*

Connie Logg & Les Cottrell

Stanford Linear Accelerator Center

Goals of this talk:

- Show you that proactive network monitoring and performance analysis is a manageable job
- Give you insight on one way it has been done
- Provide you some ideas on how to do it yourself

URL to view on-line reports:
http://www.slac.stanford.edu/netdoc/perf-rep.html

Connie Logg
email: cal@slac.stanford.edu
phone: 415-926-2879

Les Cottrell
email: cottrell@slac.stanford.edu
phone: 415-926-2523

* Work supported by the Department of Energy Contract # DE-AC03-76SF00515

Will cover

- Nature of the SLAC Network
- Goals of Monitoring
- Unix Based Data Collection and Analysis Tools
- The Data Collection Process
- The Data Analysis Process
- The Data Reduction Process
- The Data Presentation
- The Use of the Presentation for Monitoring and Analysis
- Hints for Troubleshooting Network Performance Problems
- Scaling
- Summary & Conclusion

Web pioneer White joins 'eccentric and influential' at Webby Awards

BY KYLE JAROS

David Bowie, rock icon. Francis Ford Coppola, director. Susan Sarandon, movie star. Bebo White, Stanford Linear Accelerator Center (SLAC) computing specialist.

One might wonder what White has in common with this flashy ensemble. In late April, he was elected to the International Academy of Digital Arts and Sciences, becoming one of the newest of its 350 members. Since 1998, the academy has been striving to become for the still-nascent Internet medium what the Academy of Motion Picture Arts and Sciences is for film by awarding annual Webby Awards to single out the best of the web. In Silicon Valley, it is Webbies – rather than Oscars – that count.

"I think the awards should support people who are effectively and appropriately using the medium," said White, who joined other academy members to cast his online ballot for 2002's best websites in 30 categories. He was scheduled to join an eccentric and influential audience Tuesday at the Webby Awards presentation in San Francisco.

White probably feels more at home backstage, where he helped bring to life the production now known as the World Wide Web. An information specialist who spent the last two decades addressing the computing challenges of the SLAC physics community, he was present for and closely involved in the web's inception, contributing as a technical expert, author, teacher and early advocate.

Bebo White

Joining SLAC in 1981, White handled database and network technology through the 1980s and early 1990s. He spent months at a time collaborating abroad on computing projects with Geneva's European Organization for Nuclear Research (CERN). It was there that a small group of computing experts, including Tim Berners-Lee and Robert Cailliau, developed in the late 1980s the hypertext transfer protocol (HTTP) that would become the workhorse of the modern web. It first served communication needs for physics collaborations, but soon catalyzed a monumental change in the way the Internet worked by allowing document retrieval from a web of linked computers.

"Nobody realized the web was going to take over the Internet the way it did," White said. "The first reaction was: 'That's a cute toy, a computer science exercise.'"

> **ON THE WEB**
>
> *For related information on this topic, see*
>
> http://www.slac.stanford.edu/
>
> http://www.stanford.edu/news/report/news/december12/webtams10-1212.html

White helped to bring web technology to SLAC and establish the first U.S. website in 1991; it proved a prototype for modern websites. His team linked a computer network to a SLAC database called SPIRES, for Stanford Public Information Retrieval System. "At CERN you were talking mostly about web access to static pages," White said. "SPIRES was about web access to databases. Most web-based e-commerce, digital libraries and dynamic webpages work because the web can talk to databases, and SPIRES was one of the first."

White later grew involved with the dissemination of web technology by working with the International World Wide Web Conference (IW^{3}C) series, whose Budapest conference he will co-chair next year, and Silicon Valley's SIGWEB, an association of people, labs and institutions working to advance web technology. Between projects, in 1994 White published one of the first books on hypertext markup language (HTML), a now universal webpage programming language. "The most exciting thing was how rapidly [the web] changed and how excited everybody was," White recalled. "It was literally like being in the eye of the storm."

The center of web innovation was the right place for White. "Bebo is young at heart, keen to pick up new things and try for the best," his colleague Cailliau said from CERN. "He sees merit even when it hides." According to Cailliau, this excitability helped White, an "irresistible pusher of the web," propel the technology toward the public.

Having now spent years on the web's technical side, White is enthralled to see the web recognized as a creative medium and to see how masterfully its technology is manipulated today. "The web is no more 'paper online' than TV is 'radio with a picture,'" White said. "When you write for the web you are writing for a very different medium."

For White it seems appropriate that the web, developed by a small, laboratory-based team at CERN, has retained a democratic character despite the digital arrival of big media and industry. "You go to a website and you don't know if it's a multimillion-dollar corporation or a 13-year-old kid working from his garage. Both can control the medium," he said.

Seeing the web ablaze with commerce and ideas, White finds his professional life, like that of millions of others, has become increasingly centered around it. At the same time, the web has remained a window to physics, the field in which he holds his master's degree. "At SLAC, I can do computing and still be close to and understand the science going on," White said. "I feel fortunate that I still love it and still learn stuff." SR

Kyle Jaros is a science writing intern at SLAC.

111th Commencement

Graduate School of Business students shatter gift record with $300,000 pledge

BY HELEN CHANG

In a ceremony Saturday at Frost Amphitheater, 413 candidates for graduate and doctoral degrees from the Stanford Graduate School of Business accepted the congratulatory handshake of Dean Robert L. Joss in the school's 77th commencement. This year's graduating MBA Class of 2002 received special recognition for its collective record-breaking gift to the Business School.

The diploma ceremony honored the 440 people awarded degrees this academic year: 374 received MBAs, six of them joint JD degrees from Stanford's School of Law; nine received doctorates; nine received the degree of Master of Arts in Business Research, which recognizes the accomplishments of students who enter but do not complete the doctoral program, yet complete work equivalent to a master's degree; and 48 graduates of the Stanford Sloan Program received the Master of Science in Management degree.

"This year's class is extraordinary for several reasons," said Dan Rudolph, senior associate dean, noting that a record 88 percent of the Class of 2002 participated in the gift, and that the $300,000 pledged "shattered the old dollar record by more than $130,000."

> **'You came to us with enormous talents and experiences; you leave enriched in experiences and in knowledge,' Kreps said.**

The gift was made during a very difficult economic environment. "And the old model of picking a very specific building project has been replaced this year by a much more general grant to fund several student fellowships and expand the second-year seminars," he said.

Of the MBA graduates, 38 learned that they had achieved distinction as Arjay Miller Scholars. Named in honor of the fourth dean of the Business School who was on hand personally to congratulate them, Arjay Miller Scholars are ranked in the top 10 percent of their class by academic performance.

At the top of the class, Robert Scott Berg was named the Henry Ford II Scholar and presented with a check for $15,000. Judged by classmates as having contributed most to the fulfillment of the goals of the Business School in his active participation, initiative, leadership and personal integrity, Damon A. Vangelis received the Ernest C. Arbuckle Award.

"You came to us with enormous talents and experiences; you leave enriched in experiences and in knowledge," said David Kreps, senior associate dean for academic affairs. "But this is not goodbye. We hope and expect that you will remain a member of our community, now with the new title of alumnus, but engaged in the same basic activity of growing personally and intellectually, giving to the community as you take from it." SR

Russian teen prodigies earn third set of master's degrees; Next step, doctorates

Russian sisters Anjela and Diana Kniazeva prefer econometrics to hiphop and CNN to MTV, but they insist they're just plain teenagers.

"We're not seriously different," says Diana, who just turned 15.

However, their considerable academic achievements belie that claim.

On Sunday, the Kniazevas earned their third master's degrees – this time in international policy studies. They already hold graduate degrees in law and economics, plus two bachelor's degrees from Russian universities. At the ages of 10 and 11, the sisters graduated from high school in Moscow with straight-A grades.

"They are true prodigies," says history Professor Norman Naimark, director of the International Policy Studies (IPS) Program. "The girls are extremely talented, and I'm glad that IPS was a part of their education. I treated them like the other master's students. Academically, everything went splendidly."

The sisters came to Stanford on a scholarship funded by the Russian president that allowed them to study anywhere for one year. On campus, the teens took courses alongside mid-career professionals on subjects ranging from quantitative methods to international economics.

The teens already speak English, French and German, so at Stanford they took up Spanish and Italian. "It's been great at Stanford," said Diana. "The best financial and economic education is [available] here. In Russia, there are very few elective courses. Here, it's much more flexible."

Outside the classroom, the Kniazevas spent time in-line skating, biking and listening to music – all in the company of their parents, Yulia Kniazeva and Yevgeny Bykov. Naimark said the parents accompanied their daughters to Stanford because they are minors and, as a result, the girls had little interaction with other students. "Socially, the parents are very, very involved," Naimark said. From an early age, the parents have hovered over their brainy offspring and guided their achievements.

Despite hitting the books much of the time, 16-year-old Anjela said she enjoyed her time on the Farm. "The diversity of the community at Stanford taught us about the dynamic interchange of cultures," she said in studied English. In a more relaxed tone, Anjela admitted that the balmy California climate sometimes made it hard to study. "It's like summer all around the year here," she said.

Both girls are unfazed by the wide age gap with their academic peers. In fact, Anjela said she is thrilled to have done so much at such a tender age. "I feel like I'm happier than ever," she said. "The time I invested when I was younger is paying off."

The sisters want to become successful businesswomen or researchers when they grow up. They are interested in transitional economies and, on campus, assisted Hoover Senior Fellow Alvin Rabushka, an expert on Russia's flat tax. Rabushka said he asked the Kniazevas to help him translate Russian tax forms. "It was incredibly exhilarating to see brilliance," he said. "After two hours with these teenage girls, every ounce of mental energy in my body was drained."

The teenagers' next step on the academic ladder is to earn doctorates. They are weighing offers from five American universities, including Columbia and New York University. They also may return to Russia for some time for family reasons. "I miss grandma," said Diana. SR

eworld
Living with tech
eworld@usatoday

eBriefing

The news behind the Net

By Janet Kornblum

First U.S. Web page went up 10 years ago

It's hard to imagine what might have happened with the Web if Paul Kunz had skipped a meeting in Switzerland 10 years ago.

Wednesday marks the 10th anniversary of the first U.S. Web page, created by Kunz, a physicist at the Stanford Linear Accelerator Center (SLAC). He says that if World Wide Web creator Tim Berners-Lee hadn't insisted on the meeting, the Web wouldn't have taken off when it did — maybe not at all.

Kunz had heard about Berners-Lee's Web project, but frankly, "I wasn't very interested," he says. After all, the Internet and e-mail were already standard among scientists. The Web made it possible to graphically link to documents on other computers, but it was hard to imagine the implications.

By Linda A. Cicero, Stanford News Service

Kunz: First Web page gave access to physics papers.

Kunz, who was meeting with various scientists at CERN, the European Organization for Nuclear Research, grudgingly agreed to a 3 p.m. meeting.

By 6, Kunz was sold on the Web. The two scientists linked a computer near Geneva to one at SLAC. It was the first time that the Web was on the Internet.

Kunz went home and created what was to become the first Web page on a U.S. computer; it gave scientists easy access to SLAC's database of physics papers.

The page went up at 4 p.m. on Dec. 12, 1991. A month later at a conference in France, Berners-Lee clicked over to Kunz's Web page and searched the database. The scientists were sold.

"It was a very dramatic moment," Kunz says. "I realized without that last piece in the demo people would have forgotten about the Web before they got home." Instead, they went home and told all their colleagues. Then they started creating their own pages, and the rest, as they say, is history.

Kunz, who turns 61 this month, is reflective. He knows Berners-Lee might have found another way to sell his invention, but he also might have given up. "He had already spent about a year trying to get CERN people interested with little success," Kunz says. "In hindsight, it seems what I did was very important. It might be that it saved the Web."

- ▸ **webcast.cern.ch/Projects/WebLectureArchive/kunz**
- ▸ **slac.stanford.edu/history/announce/EarlyWebAnnounce.html**

E-mail jkornblum@usatoday.com

'Photoshopper

By Jefferson Graham
USA TODAY

Hundreds of people stop by Drew Curtis' site, fark.com, every day to practice and view the latest examples of a unique, homegrown Net art.

It's not an "art" in the traditional sense that its creators need discernible skill. To practice it, you don't have to possess drawing or painting talent — just access to photo-imaging software, in particular Adobe's industry-standard Photoshop, and enough idle time to tinker with photos by adding thought balloons, bodies or faces or, if you're feeling really inspired, remaking movie posters to suit the theme of the day.

You may not have heard of "Photoshopping," but you've certainly seen some of its more notable examples since Sept. 11 — passed from person to person via e-mail or posted on the Web: Osama bin Laden with Bert from *Sesame Street*, perhaps, or a clueless tourist on the observation deck of the World Trade Center as the airplane approaches.

But this Net folk art of doctoring images for fun and amusement extends well beyond the borders of the USA and the events of recent months. Helping fuel the trend: the popularity of sharing photos on the Web, the general acceptance of the Internet as an entertainment medium, and the booming sales of digital cameras, many models of which come bundled with imaging software that makes manipulation easy.

"Parody has always been popular on the Web," says Daniel Kurtzman, who runs the political humor area of About.com. "But the war on terrorism has created a cottage industry, with Photoshopping growing astronomically."

Perth, Australia, sound engineer Cheyne Conrad, 30, says she Photoshops for "stress relief" and considers it a natural extension of making collages, which she has always enjoyed. "It's also an outlet for social comment and political satire, and the message is conveyed immediately."

Fark inspires visitors by displaying ordinary photographs — from skateboarders to Santas — and asking users to have fun with them. "I'm constantly amazed by the stuff they come up with," says Curtis, whose day job is running an Internet provider in Lexington, Ky.

Traffic to Curtis' site is on the upswing, too: Jupiter Media Metrix reports that fark.com drew 227,000 individual users who visited once or more in October. (Fark's Photoshop links are easy for newcomers to miss; the icons often are buried among its collection of weird news from around the Net, and you have to click on the "comments" to view Photoshoppers' work.)

Other showcase sites include B3TA and Worth 1000, which will hold contests for picture of the day when it launches Jan. 1. Most Photoshopping tends to the humorous (at least in the eyes of its creators), and sites that collect such work often skirt the edge

Creative outle
President Bush as "The Turbanator," from About.com, and the Statue of Liberty as depicted by Bob Borries.

La rassegna stampa di Epic*link*

Mercoledì 5 dicembre 2001

TESTATA	PAG	TITOLO ARTICOLO - Abstract
News sui nostri azionisti		
24 ORE	14	CAVI PIRELLI PER LA RETE OTTICA IN TRENTINO Ai servizi in banda larga che saranno forniti sulla rete in fibra ottica potranno accedere tutti i clienti del Cedis (Consorzio elettrico di Storo), un parco utenti costituito da famiglie (60%), imprese artigiane, industrie e aziende del terziario.
NEW ECONOMY	1	LA LARGA BANDA HA BISOGNO DI CONCORRENZA Prosegue il dibattito sulla banda larga. Necessario promuovere la competizione a livello regionale. La richiesta dei servizi si modifica nel tempo, sia sotto il profilo tecnologico, sia sotto quello dei fabbisogni della clientela.
NEW ECONOMY	2	LA MARCIA DELLA FIBRA OTTICA Sfruttando le infrastrutture elettriche saranno cablati alcuni comuni del Trentino. Il progetto nella Valle del Chiese è promosso dal Consorzio di Storo e dall Pirelli.
Broadband		
NEW ECONOMY	2	I TANTI NODI DELLA LARGA BANDA IN ITALIA Nella banda larga l'Italia è il fanalino di coda.La situazione attuale a confronto tra lo scenario americano e le future strategie per la connettività.

f i ® s † m ¤ ñ d @ ¥
PEER-REVIEWED JOURNAL ON THE INTERNET

FM Interviews

Louise Addis

I was so excited that I almost dropped the book.

I was reading *Weaving the* ***Web****: The Original Design and Ultimate Destiny of the World Wide* ***Web*** *by Its Inventor*, by Tim Berners-Lee. The paragraph that caused such a reaction described how **Paul Kunz**, a California physicist, had returned from a visit to CERN (Conseil Européen pour la Recherche Nucléaire; now officially named the Organisation européenne pour la recherche nucléaire or European Organization for Nuclear Research) in Switzerland where he had seen an interesting technology developed by Berners-Lee. **Kunz** came home to the Stanford Linear Accelerator Center (SLAC) and told Louise Addis, a SLAC librarian, about this new invention, the World Wide **Web**.

Berners-Lee states that Louise saw the **Web** "as a godsend - a way to make SLAC's substantial internal catalogue of online documents available to physicists worldwide." [1] He further states that "under Louise's encouragement SLAC started the first **Web** server outside of Europe." [2]

The first **Web** server created in the United States - set up anywhere outside of its birthplace at CERN - was established at the encouragement of a librarian!

As someone working in both the **Web** world and the library community, I decided then that I had to meet, or at least talk to, Louise Addis. I knew that I could learn something from the librarian who saw the potential in this revolutionary technology before the majority of the world had even heard of the Internet.

I tracked Louise down using the **Web** (of course!) and e-mail. When we first spoke on the phone, I told her how excited I was to meet the person who was responsible for setting up the first **Web** server in the U.S. Louise laughed and disclaimed total responsibility for bringing the **Web** to America. She told me about her colleagues, **Paul Kunz** and Tony Johnson and others, who were critical to the launch of the first **Web** server at SLAC. At one point in the conversation, she noted that the history of the **Web** technology is rather like *Rashomon*, Akira Kurosawa's 1951 movie in which the same story is told from widely diverging perspectives. It appears to be true that there are a number of stories about how the **Web** came to America. This is one of them. - Melissa Henderson.

First Monday (FM): Could you first tell me a bit about your academic background? Where did you study for your undergraduate degree?

Louise Addis (LA): I studied International Relations at Stanford University. I always had an interest in the sciences, however.

FM: And from Stanford?

LA: I initially went to work for a publisher of the "Dick and Jane series" in the claims department and as a sales correspondent. I then moved to the Stanford Research Institute Radio Science Lab where part of my job was handling classified documents. I was thinking about going to library school when a colleague encouraged me to apply for a position in the soon-to-be-opened library for Project M.

FM: Project M?

LA: "Project Monster" now known as the Stanford Linear Accelerator Center. At that time, it was still on the drawing board and had no funding. We started our library in a warehouse on the Stanford campus. It was the hottest place on campus; they ran water on roof to make it more bearable!

FM: How long were you at SLAC?

LA: From before the beginning! It was very exciting to build a library. I started as the assistant to the librarian in 1962 and retired in 1994 as Associate Head Librarian. Since retirement, I have continued to work off and on for the SPIRES-HEP database project at SLAC and most recently on Y2K conversion.

FM: What were you working on in your early years at SLAC?

LA: From about 1969 on, I was working with the SPIRES-HEP database of particle physics documents. SPIRES (initially, Stanford Physics Information Retrieval System, then Stanford Public Information Retrieval System) database management system was the product of brilliant programmers at Stanford. SLAC was a guinea pig for SPIRES, which was originally designed for an IBM mainframe environment. SLAC sponsored development of the Unix version, which is now in use.

FM: Did anything in your education or background prepare you for the higher level computing work you were doing with SPIRES-HEP or the **Web**?

LA: I got into programming and database development by picking up information on my own. If I needed to know something, I asked someone to show me how to do a particular task. Then I went back to the Library and tried it on my own.

When the SLAC Library was started, the Stanford University computing center was a little building with keypunch machines and a line printer. At the Stanford Research Institute, I had seen an "advanced serials handling system" which meant that each serial had its own punch card. This was cutting edge! So when we built the Library at SLAC, we put the serials on punch cards. I remember walking across campus with my stack of punch cards. So advanced!

FM: What was so important about SPIRES-HEP database of particle physics documents?

LA: There weren't any library database management tools available at the time and SLAC needed something to meet the unique needs of particle physics researchers. Research papers produced by the particle physics community often have hundreds of authors. The record is 1,200 authors! We were committed to listing all authors or researchers and needed a tool that would accommodate us.

Research papers produced by the particle physics community often have hundreds of authors. The record is 1,200 authors.

The SPIRES-HEP database covers the full literature of particle physics and is a collaborative project between a number of particle physics labs, primarily SLAC and DESY in Hamburg. It's so very flexible, so it's constantly developing. We've always been able to easily manipulate the database to do whatever we needed to do.

Of course, there are two sides to any database - the technology side and the content side. The content side is really what ultimately determines the value of any database. Bob Gex, then chief librarian at SLAC, was responsible for the content side, along with our long-time and remarkable preprint librarian, Rita Taylor. And from the beginning, there have been major contributions of content from our DESY partners.

FM: When did the **Web** first arrive at SLAC?

LA: On December 12, 1991, the first **Web** server in the U.S. was established at SLAC to help improve access to the SPIRES-HEP database. For a number of years, SLAC had been providing remote access to the SPIRES-HEP database. In fact, by the time we launched our first **Web** server, the SPIRES-HEP database had nearly 5,000 registered users in 40 countries.

George Crane had developed a remote interface, initially using BITNET. Researchers could send a search query to SPIRES-HEP via an interactive messaging tool, which was sort of like the instant messaging tools of today. Or they could request information via e-mail. When Internet access came along, the interactive messaging was no longer available; researchers had to use e-mail. Researchers were less satisfied with the e-mail interface; they liked the instant response system better.

On December 12, 1991, the first Web server in the U.S. was established at SLAC to help improve access to the SPIRES-HEP database.

FM: Was this a motivating issue when you first thought of setting up a **Web** server?

LA: Well, the real motivating issue was a project that we been working on earlier. In 1989-1990, I was helping set up a library at the Superconducting Supercollider (SSC) project in Texas. Pat Kreitz, the SCC librarian, wanted to use the SPIRES-HEP database and wanted to present it to the SSC community in a seamless way within a graphical user interface. SSC had funded the development of an X-Windows interface to do this, but as soon as I saw the **Web** in action, I knew that it was a fast, cheap solution.

I can still remember the day **Paul Kunz** appeared in my office after his trip to CERN. He showed me the **Web** and we started moving right away. I scurried around and got some accounts that were needed. **Paul** and Terry Hung set up the server on our mainframe. George Crane was able to easily extend remote SPIRES to talk to the new interface. And I was able to make the SPIRES-HEP database write HTML.

I've always felt deeply indebted to **Paul** for spotting the potential of the **Web** for our situation and for helping develop the first **Web** server at SLAC, which proved that potential. potential.

FM: This sounds like it was really a team effort.

LA: It was! While I was working on this project - and for most of my other projects at SLAC - I was constantly seeking input and advice from colleagues. There was almost always someone around who understood the most arcane, technical issues and was willing to help or to give advice. Sometimes conflicting advice!

FM: Did this teamwork model work?

LA: Yes, it was the only way any of these projects could continue. None of us got raises or payment for this work. We did some of it on our own time. We knew that there was value in the work - and it was fun! (I hope my colleagues would agree.)

By February 1992, I had twisted enough arms to start an ad hoc **Web** development and support group, the WWWizards. The initial group was made up of Mark Barnett, George Crane, Tony Johnson, Joan Winters, Bebo White, all of whom brought important skills to the project.