#### SLAC MEMORANDUM

**TO:** SLAC Authors

FROM: Kathryn Henniss, Manager, Technical Publications Department

**RE:** Changes in SLAC Printing and Mailing Practices in 1999

**CC:** Pat Kreitz, David Leith

As Internet and Web access becomes more widespread and better access and notification mechanisms are developed, the time has come for SLAC to join a growing number of labs in replacing large-scale printing and mailing of (unsolicited) paper versions of SLAC publications with speedy online delivery systems and multimodal notification mechanisms. This memo is to keep SLAC authors informed of the nature of recent changes made in this area.

Over the last two years, the size of SLAC's online inventory of published research documents has dramatically improved:

- At present, over 4,300 SLAC Pubs, Reports, and AP Notes (representing about half of the total
  corpus) have been made available on the Web, and more are being added to the collection
  every day.
- Most new SLAC publications are now published on the Web within a day of submission, and requests for online versions of legacy documents are typically processed within a day or two.
- The goal of digitizing SLAC's entire published research record is likely to be realized within the next calendar year.

In addition to making more documents available online at faster rates, TechPubs has developed a number of electronic mechanisms to let researchers know what has been published at SLAC, and how to get to it. Additional improvements in this area are planned (see details, next page).

At the same time that these services have been improving at SLAC, more and more institutions are asking to be taken off of SLAC's hardcopy distribution lists, since they can now access what they need via the Web. The trend is very definitely towards eliminating hardcopy distribution.

Improvements in online access combined with a decline in the number of institutions wishing to receive hard copy of SLAC's publications thus underlie the decision to eliminate bulk hardcopy distribution of SLAC Pubs, AP Notes, and Technical Notes, and to reduce the size of default print runs for SLAC Reports as feasible. The savings to the Lab in annual printing and mailing costs is considerable.

Letters have been sent to all of the individuals and institutions on SLAC's current mailing lists, notifying them of the planned printing reduction. Instead of copies of every new SLAC publication, those on the current mailing lists may choose to receive monthly (hard copy) lists of newly released SLAC publications. (This information is available on the Web as well, with links to abstracts as well as the full text of the papers; see URLs on next page.) Hard copies of SLAC publications will continue to be provided on request, of course.

We are confident that these changes will achieve significant savings to the Laboratory without any reduction in access to SLAC research. In fact, Technical Information Services is working on developing better ways to keep researchers at SLAC and elsewhere informed about SLAC's research program.

The items mentioned on the next page provide an overview of what access mechanisms are available now, and what's planned in the near future. As always, comments and suggestions are welcome.

K Henriss

#### Providing better access to SLAC publications

#### Current Access Mechanisms

Several new mechanisms have been developed in the last year for improving access to SLAC's publications. These mechanisms are available to the world-wide community of researchers interested in SLAC publications, as well as to SLAC authors locally.

Resource	URL	Description	
FastFind	http://www.slac.stanford.edu/pubs/fastfind.html	Convenient search form for find- ing SLAC publications available online.	
This Month's SLAC Publications	http://www.slac.stanford.edu/pubs/thismonth.html	List of titles (with URLs) released each week/month. Updated daily.	
TechPubs Preprints Mailing lists	http://www.slac.stanford.edu/pubs/notification.html	e-mail lists for weekly/monthly notification of newly released SLAC publications.	

#### In Development...

TechPubs and the SLAC Library are working on developing a more sophisticated electronic mailing list system which will allow users to receive (periodically, or as-released) e-mail notification of documents according to author-selected topical distribution categories:

- A (Accelerators)
- E (Experimental)
- I (Instrumentation)
- T (Theory)

Previously, these author-selected distribution categories were used to determine the number of printed copies, as well as to target the recipients for a given publication. Thus, individuals or institutions who were interested only in some subject areas would receive only papers within those distributions. Compound categorizations of papers were not uncommon (e.g., T/E, E/I, A/E/I, etc.).

In the system planned for the near future, we will duplicate this traditional print-based functionality electronically by providing e-mail notification of newly released documents to anyone who wants it. The option of receiving announcements of all new papers, or only those in selected subject categories will then be available to researchers at SLAC and elsewhere. The only change for SLAC authors in terms of registering titles and getting SLAC document numbers is that the traditional distribution categories will eventually relate to *e-mail* rather than *hard copy* distributions.\*

<sup>\* &</sup>quot;N"(None) and "M" (Minimum) distributions which relate only to print quantity, will likely be eliminated.

#### STANFORD LINEAR ACCELERATOR CENTER

TO:

Distribution list for SLAC publications

FROM:

Kathryn Henniss

Manager, SLAC Technical Publications Department

Stanford Linear Accelerator Center, MS 68

Menlo Park, California 94025

RE:

Notification of changes in SLAC printing and mailing policy in 1999

#### Dear Colleagues,

Beginning March 1, 1999 SLAC will adopt a practice now followed by an increasing number of DOE laboratories and other research institutions in discontinuing its printing and mailing of unsolicited copies of SLAC publications. After March 1, paper versions of individual publications will be mailed out on a perrequest basis only.

This change in distribution practices should not, however, reduce access to SLAC research, as it will be offset by a number of significant improvements in Internet access to SLAC publications made in the last two years. Those improvements are detailed in this memo.

Monthly notices of all newly released SLAC publications will be sent by regular mail to individuals on the current distribution list who respond by returning the form attached to this letter. Non-respondents will be removed from SLAC's distribution list, although you should be aware that SLAC's publications are still accessible to everyone via the World Wide web.

Electronic resources for accessing SLAC publications now include the following:

- A Web-based list of just-released publications, with links to the electronic full text, is now maintained on the SLAC Web site. This list is updated daily, and provides the quickest way to stay informed of new publications from SLAC. (http://www.slac.stanford.edu/pubs/thismonth.html)
- PDF and/or PostScript versions of all recently released SLAC Pubs, SLAC Reports, SLAC WPs and SLAC/AP Notes are available on the SLAC Web and are linked to entries in the SPIRES-HEP database. An easy-to-use search form on the Web has been developed to facilitate access to these online versions of SLAC publications. (http://www.slac.stanford.edu/pubs/fastfind.html)
- Electronic mailing lists have been created to which anyone may subscribe that provide comprehensive listings of newly released SLAC publications. Users may choose to receive this notification weekly or monthly. (http://www.slac.stanford.edu/pubs/notification.html)

In addition, legacy SLAC publications are now processed for Web access upon request. If an older SLAC

Pub, Report, or AP Note you are looking for is not already available on the SLAC site (over 4,300 publications are presently available), the Technical Publications Department will process it for you on demand (a Web form is available at http://www.slac.stanford.edu/grp/techpubs/postreq.html).

We are confident that the new mechanisms described above will provide you with more timely and costeffective access to SLAC's published research than has been possible in the past when distribution was limited to mailing thousands of paper copies. Instructions for receiving e-mail or paper notification of newly released SLAC publications and requesting hard copy versions of particular publications may be found on the form accompanying this letter.

If you use a mail application like Eudora, URLs contained in the regular mail messages sent to these mailing list will provide you with direct links to the Web version of each publication.

#### To receive regular notices of newly released SLAC publications...

You may choose to receive paper and/or e-mail notices of SLAC's newly released publications. Paper notices will be sent out monthly to individuals who respond to this letter by March 31, 1999. E-mail notification is sent to members of the TechPubs Preprints Mailing lists at the end of each business week and the end of every month. The listings will include the title, authors, and URL of the full-text and abstract of each publication, as well as an indication of the subject area (e.g., Accelerators, Theory, Experiments, etc.)

<b>&gt;&lt;</b> _	Clip and mail this form.			
	Check here if you wish to receive monthly hard copy notices of newly released SLAC publications. Please update your mailing address on the label below, if necessary:			
	(mailing label goes here)			
	Mail this form to:  SLAC Library, Attn. Distlist Coordinator			
<b>×</b> .	Stanford Linear Accelerator Center, MS 82 2575 Sand Hill Road Menlo Park, CA 94025			

#### To receive e-mail notification

To subscribe to the TechPubs Preprints Mailing lists, send an e-mail message to majordomo@slac.stanford.edu, and include one or both of the following in the body of your message:

```
subscribe preprintsweekly-1 Note: "-l" is a lowercase L, as in "list." subscribe preprintsmonthly-1
```

Shortly after you send your message, you should receive an e-mail confirming that you have been added to the list(s) and information on how to unsubscribe. This information is also available on the Web at:

http://www.slac.stanford.edu/pubs/notification.html

#### To get paper copies of SLAC publications after March 1, 1999

Hard copy versions of SLAC publications will be sent out *by request only* after March 1, 1999. There are two ways to request a paper version of a SLAC publication:

1. **By regular mail**: Return a copy of the weekly or monthly notice you will have received (by regular mail or e-mail), indicating with check marks the publication(s) you require. Requests should be sent to the address below, and will be processed as they are received:

Printing and Distribution Coordinator SLAC, MS 68 2575 Sand Hill Road Menlo Park, CA 94025

2. **By e-mail**: If you know which document you loozwking for, send an e-mail request to crystal@slac.stanford.edu, including in your email the document number (e.g., slac-pub-7770), the document title, and the first author's name. Email requests for individual publications are processed as they are received.

The second of th

STAC ARCHIVES COLL

SEPIES 2 SHALL SHEET SHE

Document: http://slacvm.slac.stanford.edu:80/FIND/slacinst.html



#### Introduction to SLAC

SLAC 8 Dec 1993

This panel is under construction.

The Stanford Linear Accelerator Center (SLAC) is a national laboratory operated by Stanford University for the US Department of Energy. SLAC has been in continuous use for over 25 years in a national research program that has made major contributions to our understanding of nature. The Center is one of a handful of laboratories worldwide that stands at the forefront of research into the basic constituents of matter and the forces that act between them.

SLAC does experimental and theoretical research in elementary particle physics using electron beams, plus a broad program of research in atomic and solid state physics, chemistry, biology and medicine using synchrotron radiation. Scientists from all parts of the United States and throughout the world participate in this work. There are active programs in the development of accelerators and detectors for high energy physics research and of new sources and instrumentation for synchrotron radiation research.

SLAC was founded in 1962, and the Stanford Synchrotron Radiation Laboratory (SSRL) came into being in 1979 as a national users' facility. SSRL became part of the SLAC facility in 1992. Their combined staff is currently about 1400, of whom 150 are Ph.D. physicists. At any given time, there are typically 300-400 physicists from other institutions participating in the high energy physics program and 600 scientists in the synchrotron radiation program.

SLAC has the following major facilities:

- The Linac, a three-kilometer (or two-mile) long linear accelerator, capable of producing electron and positron beams with energies up to 50 GeV
- SPEAR, a storage ring 80 meters in diameter now used as a synchrotron radiation source
- PEP, a 30 GeV colliding—beam storage ring, 800 meters in diameter
- SLC, a 100 GeV electron-positron linear collider
- Several large particle-detection facilities



as shown in this gerial view



Keyword:

Go Back | Previous

Next · Save...

Search...

· Clone

Poetscript Nevigate Mttp://www.slac.stanford.edu/winters/pub/www/kgrisl.gif Documents Manuals

Go Back Previous Next · Save... Search... · Clone Close Window Keyword:

The state of the s

STAC ARCHIVES COL

SERIES SUBSERIES FOLDER

AlWWizards Mfg 10-5-94 A Henders: John Halperin, Bebo White, Tony Johnson Les Coltrell, Pat Kreitz, Joan Winters

Agenda: / Dickens Memo-final cul-/ CERN VM Server \* Màc Server Mail Lists Home Pages Pat Kreitz Bebo Pat Clancey Les Coltrell

\* Defened to Clancey's rickness

Dickens Memo-final cut
Focussed on time estimate for phases.

Bebo agreed to get CERN VM server installed by end of Thursday Mail Lists / Net news Groups

(2)

www-l: for de ign people (former wizards) to discuss design issues/status

www-tech: tech committee (subset of www-e In neetings etc.

www-admin: problem reporting, equivalent to

reporting to helpa slac. stanford.edu

All ahove in major do mo 2 access va <namexos lac. stanfod.

slac users www news group (may exist already)
Need publicity etc.

nuvertes of www Teal grap Future Topics

Disclaimers

Neb Page

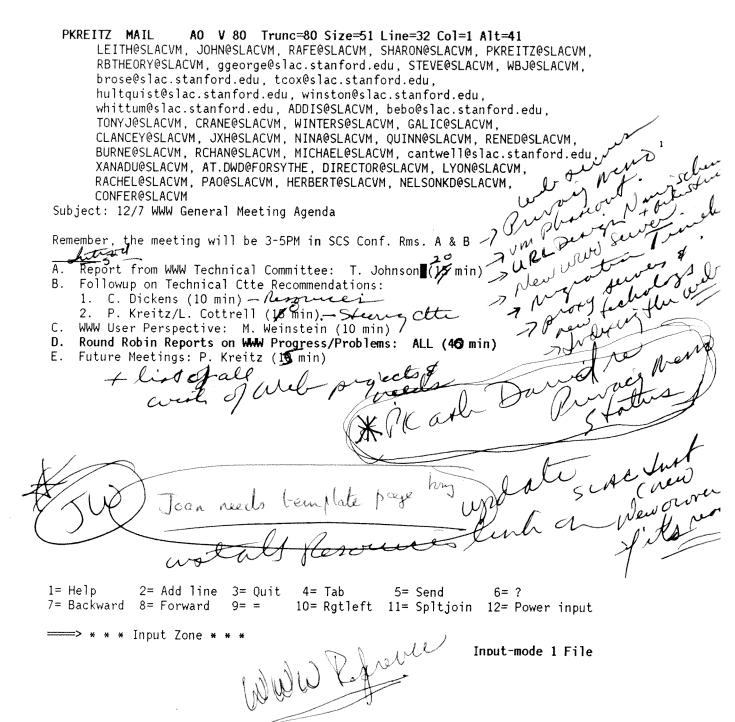
Meet with Cluck to go over memo. Netnews slac. users. www publicity

WWW hother forme Pose In cottrell/memos/www

#### General WWW Meeting Agenda:

Weds. 3:00-5:00 Dec. 7

- 1. Introduction: P. Kreitz
- 2. Report from WWW Technical Committee: T. Johnson (20 min)
- 3. Followup on Technical Ctte Recommendations:
  - a. C. Dickens (10 min) 3.30 3.40
  - b. P. Kreitz/L. Cottrell (10 min) 3:40 -3:50
- 4. WWW User Perspective: M. Weinstein (10 min) Sout Geep Civil
- 5. Round Robin Reports on WWW Progress/Problems: All (40 min)
- 6. Web Committees and future meetings: (5 min)
- 7. Brainstorm wish list of all WWW projects/needs (10 min)



Commendo ela. to Aultino hy Close of To: David Leith
From: Pat Kreitz & Les Cottrell
Subject: DRAFT: Creation of WWW Policy Working Group

During this transitional stage of WWW at SLAC, from ad hoc into a supported production service, there is a need to put together a framework for the future support for WWW at SLAC. In particular there is a need to expose the use of WWW to a wider audience, to draw in information providers from a wider SLAC audience, to provide a lasting framework to enable information users and providers to know how and where to find and provide information, to ensure the information is in line with SLAC policies and guidelines, and to formalize the support structure.

#### ACTION:

Establish a non-technical WWW Steering Committee to recommend to the AD's what further actions should be taken in the areas listed below to continue the implementation of WWW as a communication tool for SLAC business.

#### CHARGE:

- \* Create an outline of the general policies and procedures to organize and administer WWW at SLAC. Issues that need to be addressed include:
  - recommend mechanisms for selecting the kind of information to be made available, for creating a general look and feel for SLAC information, and for assigning ongoing responsibilities
  - review existing WWW recommendations, policies and practices at SLAC
     recommend initial WWW policies and how existing SLAC policies should

be applied to WWW

- recommend how to deal with issues including registration of pages, contact people, publicity, documentation, guidelines, maintenance, privacy, security, supported formats and tools, on going support level expectations, growth, proliferation of WWW servers at SLAC, advisory groups, coordination of WWW matters within SLAC and outside especially with Stanford University and the DOE
- \* Identify what needs to be done and by whom, and recommend on-going activities to provide support, advice and guidelines. More detail on the specific areas needing support is given in the WWW Technical Committee's report to Chuck Dickens on WWW VM Migration.

In the process, the Steering Committee will coordinate with other areas and groups such as the WWW Technical Committee, networking, legal, other sources and information services at SLAC and its collaborators.

#### DEADLINE:

The group should produce an interim report in 2 months since the recommendations are needed as soon as possible to be presented to the ADs.

#### MEMBERSHIP:

The group should be small (less than 8 people), energetic (members will need to be very interested in the issues and prepared to dedicate substantial time during the lifetime of the group) and include people who can represent a wide range of divisions and WWW users.

The committee should be free to establish ex officio members if needed and will invite a number of Web users and creators to work with it to meet its charge.

This committee will consult closely with the existing WWW Technical Committee; and hold information meetings with the wider community of implementers and major users to obtain input on needs and structure.

A chairperson should be chosen and appointed by the ADs.

From disk PAK191

R2.08 R208004) with BSMTP id 2009; Wed, 02 Nov 94 08:56:29 PST Received: from DIRECTORY-DAEMON by SCSUL.SLAC.STANFORD.EDU (PMDF V4.3-10 #6987) Received: from SCSUl.SLAC.STANFORD.EDU by SLACVM.SLAC.STANFORD.EDU (Mailer id <01HIZXY8MKK0000VUS(GSCSU1.SLAČ.STANFORD.EDU); Wed,

Received: from unixhub.SLAC.Stanford.EDV by SCSUI.SLAC.STANFORD.EDV (PMDF V4.3-10 #6987) id <01HIZXXTAXDC000HUZ@SCSUI.SLAC.STANFORD.EDV); Wed, 02 Nov 1994 08:54:54 -0800 (PST) 02 Nov 1994 08:55:15 -0800 (PST)

from oracle.SLAC.Stanford.EDU id AA09652; Wed, 2 Nov 94 08:53:26 PST Received: by oracle.SLAC.Stanford.EDU (AIX 3.2/UCB 5.64/SLAC 920508) Received: by unixhub.SLAC.Stanford.EDU (4.1/SLAC 920508) Wed, 2 Nov 1994 08:53:26 -0800 id AA16415;

From: Les Cottrell (COTTRELL(BUNIXHUB.SLAC.STANFORD.EDU) Date: Wed, 02 Nov 1994 08:53:26 -0800

Subject: Updated Memo for Leith

To: pkreitz@SLAC.Stanford.EDU Message-id: <9411021653.AA16415@oracle.SLAC.Stanford.EDX X-Envelope-to: pkreitz@SLACYM.SLAC.Stanford.EDU Content-transfer-encoding: 7BIT

ದ Below is the updated memo. I will pass a copy to Chuck see if we can get on his calendar.

From: Pat Kreitz & Les Cottrell

Subject: Creation of WWW Interface Committee (or WWW Steering Committee)

formative stage of WWW at SLAC there is a need to put together During this

a framework for the future support for WMW at SLAC.

what further actions should be taken in the areas listed below to continue the implementation of WWW as a site-wide communication tool. Establish an ad hoc WWW Interface Committee to recommend to the AD's

- \* Create an outline of the general policies and procedures to organize and
  - administer WWW at SLAC. Issues that need to be addressed include:
     assigning responsibilities, registration of pages, contact people, subject coverage, publicity, documentation, guidelines, maintenance, privacy, recommended formats, WWW policies
    - on going support level expectations, advisory groups,
- other areas such as SLAC publicity, networking, legal, coordination with

information and information services at SLAC and its collaborators Identify what needs to be done and by whom, and recommend on-going achivities to provide support, advice and guidelines. — ( 9 or each factor) other sources of

The group should DEADLINE:

presente 4 months) since the recommendations are tus rej terim rovide an and it show as soon as postible, have a short ۹

1 MEMBERSHIP:

need to be very interested in the issues and prepared to dedicate substantial time during the lifetime of the group) and include people who can represent: Lab management; durance public relations; expertise on legal issues pertaining to security policy; and others deemed appropriate by the ADS. intellectual property, privacy, and copyright; expertise the Library; central computing support; SLAC MAN users, The group should be

with the WWW Technical Committee; hold on information meeting with the wider "WWW Appended Committee" to obtain input on needs and structure, and consult individually with a wide spectrum of Lab staff.

Les Cottrell

Mail Stop 97, Stanford Linear Accelerator Center, P.O. B. 4349, Stanford, CA 94309

Phone (415)926-2523, FAX (415)926-3329

URL http://www.slac.stanford.edu/~cottrell/cottrell.html

--: R: -- Reply from PKREITZ

08:59 -0800 (PST) 02 Nov 1994 Date: Wed, From:

"Patricia Kreitz (415) 926-4385" COTTRELLGUNIXHUB.SLAC.STANFORD.EDU

Subject: Re: Updated Memo for Leith

excellent

10 Ba

Page

MAILSTMP FILE

after.

A1

Dated 11/02/94 08:59:28

(less than 8 people)

energetic (members will

This committee will consult closely

(PKREITZ/GSLACVM)

In-Reply-To: COTTRELL(@UNIXHUB.SLAC.STANFORD.EDU -- 11/02/94 08:56

SA TO CE

'a fipal report

creation, after its

months

and pre

TEST WorldWideWeb SLAC Home Page

Document: | http://www-slac.slac.stanford.edu:5080/slac.html

Network organizations:

BARRNet, CREN/BITNET, ESnet, HEPIC, SuperJANET.

Local area resources:

Stanford University and its School of Medicine, more local area resources.

Other federal resources:

DOE, FedWorld, the MetaCenter, NASA, NCAR, NERSC, USGS, more federal agencies.

Professional Societies:

AAS, ACM, AIP, APS (including PACS), NAS; HEPIX.

Other information sources:

academic fields (the WWW Virtual Library), GopherSpace, grab bag, hacker's jargon, LISTSERV lists, Netnews FAQs, vendors.

#### SLAC WWW Support

WWW at SLAC is supported by the WWW \_Technical\_Committee, to whom you should address questions, comments, complaints, etc. See What's New for updates to SLAC's WWW pages, Major Changes for more system-related modifications, and SLAC WWW Server Statistics for usage data. You may also find the Old SLAC Home Page and the Test\_SLAC Home\_Page, helpful.

The <u>WWW Project</u> was initiated at CERN, from which support is still coordinated. <u>WWW</u> contributors currently come from diverse parts of the world. See the WWW bibliography for initial pointers to major topics and WWW resources for pointers to SLAC authoring and other service providers' materials.

This page is intended for people experienced with WWW at SLAC ("refreshers").

This version of the SLAC Home Page was created by Joan Winters and evolved from part of the original one created by Tony Johnson and updated by various SLAC WWWizards. This version of the SLAC seal was created by Terry Anderson.

<u>Winters</u>

Keyword:	
Go Back Previous Next Save Search Clone Close Window	

<sup>\*</sup> Access to this link is restricted to SLAC users.

#### SLAC WWW Resources

SLAC 6 Dec 1994

This page is under construction.

If you want to create some Web pages or supply related services, here's some information to help you get started.

N.B. This page is very preliminary and incomplete. The information pointed to is often just an initial draft or preliminary ideas about procedures, standards, conventions, and software. Much will change. For your information now:

- Draft Template for SLAC WWW Page
- <u>Draft Summary of New WWW Server Rules</u>
- Authoring Tools for SLAC HTML Document Developers
- WorldWideWeb\_Browsers\_at\_SLAC
- Meeting\_Notes\_ of \_the\_SLAC\_WWW\_Technical\_Committee
- SLAC \_WWW \_Mailing Lists and Netnews Groups
- <u>SLAC Page-Design Playpen</u>

You may also find the following helpful: <u>WWW\_bibliography</u> <u>SLAC\_WorldWideWeb</u> <u>services</u>.

Press here to go to the SLAC Home Page.

Winters

				***************************************	······	·····
Keyword						
Go Back	Previous	Next · Save	Search	· Clone	Close Window	





Appear Street Control (Artist State) and the street State State Control (State State State

I would like to propose the formation of a user group at SLAC which addresses the issues of network information retrieval. Unless/until someone comes up with a better name, I'll call it SWUG (SLAC Web Users Group).

I would envision the following as a part of a SWUG charter:

- \* to provide a mechanism whereby SLAC WWW users and maintainers could communicate with one another re: issues of WWW support at SLAC. This could follow the Apple Support Coordinator (ASC) model.
- \* to provide a mechanism for user input on technical issues to the WWW Wizards committee.
- \* to promote the use of WWW at SLAC through relevant presentations and demonstrations; perhaps by following the original intention of the Unix Journal Club.
- \* to be a local SLAC chapter of SigWeb and provide a SLAC/HEP presence in that organization.

Involvement of users has become increasingly important given the proliferation of servers and the widespread security and privacy issues.

Of course, SWUG would be expanded to encompass tools other than WWW (esp. Gopher and WAIS) as they become available at SLAC with either the installation of clients and/or servers.

Please let me know your thoughts. In particular, when and how often should SWUG meet? I will make the necessary room arrangements.

I Integrated of SSRie Support old ter minds 2 an venent efficient Home Page (5) 3 Encyc Brillanica access 4. Web should not be only access to common dB o Classon HTML Juns/taining E Publem with egi-scripts ja graphics 7 Web Maker & Fm 216ml 8 How to get dustins of pages udded to hime page es Departments at siAe. 9. SLAC 11 P reed depts @ SLAC on welcome Page. 10. Probleach thows PC Mescie browsers 11 Jangen in SCAC HP. eg. X 500 12 holex to www SiAc resources. To: majordomo @ list. stanford. edu subscribe www-people Needle Poste Surpl' vewer In Mac Noed derection to resources. 5. Learn low to create home pages. 17 the Decumentalian of low to do scripts for 18 Apprepriation advice/ suide lines 19 It publicity/use for disemminate

1 Integrated of SSRLE Support old ter minch 2 Convenient efficient Home Page (5) ) Encyc Britanica access to web should not be only access to Classon HTML finstaathaning USERGERS Problem with cgi-scripts for graphics to 7) Web Maker Frank Training growth James How to get dustins of pages udded to hime page e.s. Departments at siAe. (9) SLAC IIP reed depts @ SLAC on welcome Page. 10. Probhad Hows PC Mesaie Knowsers (1) Jangen on SCAC HP. eg. X 500 12 holex to www SiAc resources. To: majordomo @ lest. stanford. edu subscribe isww-people 14 Norde Poste Script vewer In Mac Noed derection to resources Supplied cente diged in to create home pages. Decumentalian of how to do scripts for Apprepriaterers advice/ quide lines if publicity Juse for disemminate New Atr = 11/9/2

## Proposed URL scheme for SLAC WWW server

**DRAFT**Please send comment to Tony Johnson

In this brief memo I propose a new naming scheme to be used for documents served from the central SLAC WWW server. This memo only addresses the URL naming scheme as seen by users of the web, it does not address where the documents are actually stored.

#### Why do we need this?

Currently most of our pages have names of the form:

```
/FIND/name.html
```

The present of the word FIND is purely a holdover from the fact that our VM server was originally based on the server being used at CERN for their FIND interface. The lack of any file hierarchy makes it very difficult to split the pages into manageable chunks, and as the number of pages we are serving grows ever larger this will become more and more critical.

In idea is to change the URL's of the files once, and then publicise the new URL scheme widely. The scheme is designed to allow us to implement it on VM, and keep the same scheme as we migrate to Unix. For some time we will have to support both the old and new URL-naming scheme.

#### **URL** naming proposal

This is a proposal for the URL naming scheme for SLAC. It addresses only the portion of the name that comes after the node name. The idea is to standardize the first few levels of the heirarchy, but to allow subgroups complete control over the heirarchy below the level described here.

```
/slac.html - Home page

/ Synonym for home page

/slac/* - Top level SLAC pages (many of those currently

maintained by Joan).
```

#### Special cases/examples

```
/slac/disclaimer.html
                                 (slac disclaimer)
                                 (example of signature at bottom file)
/owner/tonyj_johnson
/slac/phonebook.html
                                (SLAC phonebook (form))
/suggestion/winters
                                (Suggestion box)
/glossary/term
                                (SlacSpeak)
                       - Spires descriptions and database accesses
/spires/*
                       - Spires freehep database (eg)
/spires/freehep/*
                      - Public information, eg description of SLAC, life at
/pubinfo/*
                         SLAC etc, stuff maintained by publications.
/exp/sld/*
                      - SLAC experiments
    /bes/*
    /babar/*
    /e145/*
/grp/scs/*
                      - SLAC groups other than experiments
   /bsd/*
    /esh/*
    /theory/*
```

#### **Other Considerations**

The inclusion of /slaconly/ in any url restricts access to slac.stanford.edu nodes.

```
<title>{Draft Template for SLAC WWW Page}</title>
 <h1>{Draft Template for SLAC WWW Page}</h1>
 <address><a HREF="/pubinfo/slacinst/slacinst.html">SLAC*</a>
 [Last modified:] {
                      5 Dec 1994 \ < /address >
 [Optional:] This page is under construction.+
 {body text, e.g.:}
 <h2>[Optional:] {Files|Information|...} at SLAC</h2>
 [Optional:]
 You may also find the following SLAC {info type} helpful: <a HREF="{URL}">{name
 { . . . } . +
 >
 [Optional:] Press <a HREF="/slac.html">here</a> to go to the SLAC Home Page.+
 <h2>[Optional:] Information Sources Elsewhere+</h2>
 {text on sources beyond SLAC}
 >
 <hr>>
 * Or link to pr-style page introducing a major SLAC group like SLD or SSRL,
   which in turn has a link to this SLAC institutional pr-style page in that
    location.
 [Optional footnote:] + Include this section only if relevant.
 [Be sure to include a footnote for any links that are restricted to
SLAC users only.]
>
[Optional acknowdedgement:] {This page was originally compiled by ....}+
<address>
[Owner:] <a HREF="/owner/winters">{Winters}</a>
</address>
[Or optional:*]
<address>
[Owners:]
<a HREF="/owner/{username1}">{username1}</a>,<a HREF="/owner/{username2}">{username2}</a>;
<a HREF="/owner/{username3}">{username3}</a>,
<a HREF="/owner/{username4}">{username4}</a>
\{\ldots\}
</address>
<hr>
where text in [] is usually omitted, e.g., [Optional:]
and information in { } is to be replaced by the author.
```

The consideration of a disch was not

SERIES SUBSERIES

SUBSERIES 202

Document: http://www.slac.stanford.edu/BF/doc/www/bfHome.html

#### **B Factory Detector Home Page**

SLAC 20 Jan 1994

#### People

[Not yet available.]

#### Information and Publications

- Calendar and Newsletter (\$BFROOT/doc/Newsletter)
- BaBar\_Netnews Groups (slac.b-factory.\*)
- On-line Documents (\$BFROOT/doc)
- On-line BaBar Notes (ftp.slac.stanford.edu in pub/bfactory/BaBarNotes)

#### **Detector Workgroups**

- Computing
- Trigger & Data Acquisition
- Physics & Simulation
- Interaction Region
- Yertex Detector
- Central Tracker
- Particle ID
- Calorimeter
- Magnet & Muon

#### Administrivia

- Collaboration Issues
- Group\_C Plan

Keyword: find title whole internet guide

Go Back

Next · Save...

Search... Clone Close Window





File Postscript Navigate Customize Documents Manuals

Help

Document: http://www.slac.stanford.edu/BF/doc/Newsletter

#### Newsletter

Up to dog
This directory area contains BaBar Newsletters. Select a newsletter based upon the date built into the filename.

#### Files

1.17.94.html 1.17.94.txt 1.24.94.html 1.24.94.txt README

Keyword: | find title whole internet guide

Go Back

Next | Save... | Search... | Clone | Close Window





R Postscript Navigate

Custombre

Documents

Manuals

活力

Document: http://www.slac.stanford.edu/BF/doc/Newslatter

# Newsletter

Up to\_doq This directory area contains BaBar based upon the date built into the Newsletters. filename. Select a newsletter

### Files

1.17.84.btml Beadwe

1.17.94.txt

1.24.94 btml

124.84.100

Keyword:

Go Back | Previous | Next | Save...

find title whole internet guide

Search... | Clone





Document: http://www.slac.stanford.edu/BF/doc/Newsletter/1.17.94.html

#### Babar Calendar And Newsletter Neek of January 17, 1994

This is the first issue of a new weekly calendar and newsletter intended to provide up-to-date information about activities of the SLAC B Factory Collaboration. The calendar section should contain future meeting announcements. The newsletter section may contain \_contributed\_ news items of general interest.

The idea and format for this newsletter borrows heavily from the similar Phenix newsletter (to whom we are indebted for the good ideas).

#### What's Inside:

Meeting Calendar

The BaBar Calendar and Newsletter

Morking in an International Collaboration

How to submit an article for this newsletter

#### Meeting Calendar

This initial issue of the calendar is undoubtedly incomplete. Please help us remedy this situation. (See below for information about reporting calendar events.)

1/18/94 [Hitlin] Group Coordinators/Steering Committee Neeting Tuesday 08:30 PST SLAC Video Conf Rm.

1/21/94 [Quarrie] Computing Subgroup Meeting (Software Architecture/Frame) Friday 13:30 PST SLAC Group C Conference Rm.

1/27/94 [Porter/Glanzman] Computing Group Meeting Thursday 09:00-16:00 PST SLAC Conf. Center Rooms CAD 2/10-12/94 [Hitlin] Collaboration Meeting

Keyword: find title whole internet guide

Close Window





Go Back

Previous

Next Save...

Search...

Clone

Title: Untitled (http://www.slac.stanford.edu/BF/doc/New

Document: http://www.slac.stanford.edu/BF/doc/Newsletter/1

Format: HTML

Show Document Source

Search... | Help | Cancel

<b>

Babar Calendar And Newsletter Week of January 17, 1994

#### HEADLINE:

This is the first issue of a new weekly calendar an intended to provide up-to-date information about ac SLAC B Factory Collaboration. The calendar section future meeting announcements. The newsletter section \_contributed\_ news items of general interest.

The idea and format for this newsletter borrows head similar Phenix newsletter (to whom we are indebted ideas).

What's Inside:

**Heeting Calendar** 

The BaBar Calendar and Newsletter

Working in an International Collaboration

How to submit an article for this newsletter

Postscript Navigate Custombe Documents Marcals

Heto

Document: http://www.slac.stanford.edu/BF/doc/Newsletter/1.17.94.txt

<PLAINTEXT>

Babar Calendar And Newsletter Week of January 17, 1994

HEADLINE:

\_contributed\_ news items of general interest. future meeting announcements. SLAC B Factory Collaboration. intended to provide up-to-date information about activities of the This is the first issue of a new weekly calendar and newsletter The calendar section should contain The newsletter section may contain

similar Phenix newsletter (to whom we are indebted for the good The idea and format for this newsletter borrows heavily from the deas

Keyword: | find title whole internet guide

Go Back

Previous Next | Save... | Search

Search... · Clone





Note Document: P Keyword: | find title whole internet guide This is a multipage Postscript document, select page: Page 2 Page Page Page 7 Page Page Page Page 1 Postscript ftp://ftp.slac.stanford.edu/pub/bfactory/BaBarNotes/BaBar109.ps Navigate Customize **Documents** Manuals 至

Go Back Previous Next · Save... Search... · Clone

File Postscript Navigate Custombe Documents Manuals

Hole

Document:

ftp://ftp.slac.stanford.edu/pub/bfactory/BaBarNotes/BaBar109.ps

#### Source Code Management System\*

Paul F. Kunz and Torry Hung Stanford Linear Accelerator Center

#### ABSTRACT

As part of the computing R&D initiated by the B Peccary Workshops, software developers at SLAC have been using source code management systems that are nodive to UNIX. Since July of 1992, the CVS system has been extrastively used and has become the system of choice by meanly all UNIX-based developers at SLAC. This ness gives some background on the choice of CVS and some guides on how to get started using the system.

#### 1. Background

Source code management systems have been available for UNIX since 1975 when the main implementation of UNIX was on a PDP-11. There are two systems that are widely known and used, both in academia and industry...

1. Source Code Control System (SCCS)[1]. This was the first system available under UNIX and is still widely used. It was developed by AT&T, who owns the linears for

Keyword:

find title whole internet quide

Go Back

Previous

Next · Save...

Search... · Clone





品 Postscript Navigate Customize Documents Manuels Hete

Document: ftp://ftp.slac.stanford.edu/pub/bfactory/BaBarNotes/BaBar119.ps

This is a multipage Postscript document, select page:

- Page 7"
- Page '6"
- Page '5".
- Page 3
- age 2
- Page 1

## Note

Once in the document you can use the Next and Previous buttons to go to adjacent

Keyword: | find title whole internet guide

Go Back

Previous | Next | Save...

Search... | Clone | Close Window





CIAC I was the set I throughout I

SERIES SUBSERIES
BOX FOLDER

Document: [http://slacvm.slac.stanford.edu:80/FIND/slac.html



#### WorldWideWeb SLAC Home Page

#### SLAC 25 Jan 1994

Use the WorldWideWeb\_ (WWW) service to gain access to a wide range of information at SLAC and elsewhere around the globe. Emphasised text like this is a hypertext link.

You may view WWW information through GUI or line-mode herougers. At least most SLAC pages have been tested on the Midas WWW X Window System browser. Note that over time links may move around on a page, migrate to others, or be removed entirely, due to the dynamic nature of the Web.

#### SLAC Information

People and organizations:

people at SLAC particle physics people and institutions.

Library:

SPIRES-HEP, Current PPF-list Books, SLACspeak slossery, other databases.

Physics Preprint Bulletin Boards (full-text postscript):

today, yesterday, last seven days, week before, that anytime

Seminary:

inday, tomorrow, this week next week anytime.

Conferences:

this month, next month, next year, next summer, all future, let me search

News:

APS What's New, SLAC Nothern SSC\_Nows.

#### SLAC Physics Program

Experiments:

BaBar, BES, mQ SLD, other.

Accelerator operations logs:

vesterday, today, this week, anytime.

T.AC Computing

Keyword: Go Back | Previous Next - Save ... Search... Clone Close Window

Document: http://slacvm.slac.stanford.edu:80/FIND/slac.html

General computing:

Amiga, Macintosh, PC, UNIX, YM HELP, YMS Help:

FreeHEP, Futures, Local Area Networking, Network Reference, Security,

SLACwide, other.

Group computing:

SCS, other.

Wide Area Networks:

BARRNet BITNET, ESnet HEPnet Internet Superjanet, other.

#### Other SLAC Information Resources

Annals Laboratory facilities, Stores\_ostalog, Telephone\_directory\_reference section, other.

#### Other Useful Information

Other institutions:

Brown, CERN, DESY, Fermilab LANL, LBL, SSC, more HEP institutions:

Stanford University (Campus and the Medical Center);

AIP (FYI and Physics News Updates), NASA, NCAR, National MetaCenter for

Computational Science and Engineering other.

Other experiments:

ALEPH DELPHI, L3, OPAL: CLEO: HERA-HI ZEUS: DQ COF: 20072 HEP.

experiments.

Other information sources:

academic fields (the WWW Virtual Library), ESnet X:500 white pages, GopherSpace grah hag hacker's jargon. LISTSERV lists. Netnews FAOs, other.

#### Support

WWW at SLAC is supported by the SLAC WWWizards, to whom you should address questions, comments, complaints, etc. See What's New for updates to SLAC's WWW pages or Major Changes for more system-related modifications. You may also find the Old SLAC Home Page or the Test\_SLAC Home Page useful.

The WWW Project was initiated at CERN, from which support is still coordinated. WWW contributors currently come from diverse parts of the world. For more information see WWW bibliography.

'This page is intended for people experienced with WWW at SLAC ("refreshers").

This version was created by Joan Winters and evolved from part of the original SLAC Home Page, created by Tony Johnson and updated by various SLAC WWWisards.

Winters

Keyward:

Go Back Previous Hext · Save... Search... · Clone Close Window





Document: | http://slacvm.slac.stanford.edu:80/FIND/slacspeak SLACSPEAK: Glossary of SLAC-related acronyms and terms

Use standard SPIRES search terms (FIND, BROWSE, SHOW, ...): Show find term CRID browse term internet search terms

Alternatively, use the WHATIS command:

find find

term

CAMAC

term SLC OR SLD

term

NEPA

whatis TCP/IP whatis NNN

Corrections/ additions to: rohan@slacvm.slac.stanford.edu

Keyword:

Go Back | Previous | Next - Save...

Search... · Clone





Correspondence (Anne Raible).

1 394

SERIES\_\_\_\_SUBSERIES\_\_\_\_
BOX\_\_\_\_FOLDER\_\_\_\_

### Memorandum

To:

Anne Raible

From:

Pat Kreitz

Subject:

ESTC software information from SLAC

Date:

12/17/94

In response to your inquiry for the software documentation for four software packages: CHEETAH, HIPPODRAW, HIPPLOPLOTAMUS, and DEPOT, I am able to supply you with more specific information about each package and, where appropriate, how interested individuals may obtain the latest information about the software.

CHEETAH: Cheetah is a data management system based on the C programming language, with support for other languages. Its main goal is to transfer data between memory and I/O streams in a general way. It was developed by Gary Word, who began working on it many years ago before he joined SLAC and left the canceled SSC to go to work for a Dallas software development company. The second author, Paul Kunz, is still working at SLAC. The software documentation and code may be obtained from anonymous ftp from ftp.slac.stanford.edu in pub/sources.

HIPPODRAW: An application to view data that is in the form of n-tuples. An n-tuple is a table of data with a fixed number of columns and a (perhaps large) number of rows. HippoDraw can create histograms and plots which are projections of the n-tuple's columns. Everything in HippoDraw is controlled via the GUI. There are no commands and no scripts. The documentation in included in the TAR file and may be obtained from anonymous ftp from ftp.slac.stanford.edu in pub/sources.

HIPPOPLOTAMUS: An n-tuple management and display package written in ANSI C with an object orientation. The management part is designed to be user friendly and also has a FORTRAN binding. Binary files use the XDR format so binary ftp can be done between machines of different architectures. Files can also be converted from or to a plain text format and from HBOOK4 format with supplied utilities. Access is from anonymous ftp from ftp.slac.stanford.edu in pub/sources.

DEPOT: Is a parts management system for the Technical Division of SLAC and contains information about the parts of the accelerator systems and equipment. It is written in SPIRES/VM and is not portable to other computing platforms. Most sites are moving off of IBM mainframes and few others have the capacity to run SPIRES/VM. At present, this system is site-specific and, in fact, needs to be replaced by a parts/equipment management database written in a more modern programming language and operable in UNIX.

Paris Assertated Marie 22. Morte Miller 1555, 20 miller Mr. Charreson x waster and

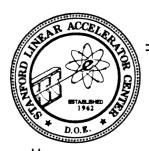
SERIES GUES COLLEGE DE LA COLL

## New Directions in Technical Publications at SLAC



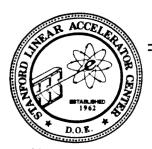
## Kathryn Henniss, Technical Publications Department Stanford Linear Accelerator Center

**September 28, 1995** 



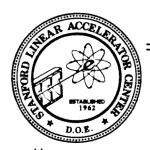
## **Today's Challenges**

- Doing more with less in light of:
  - + Increased workflow
  - Staff reductions
  - + Increased paper costs
  - Shrinking budgets
- Keeping up with changing technologies
  - + More things to learn
  - Less time to learn
- Publishing concurrently, paper and online



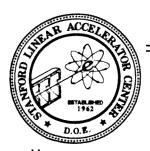
## **Strategies**

- Streamline work flow
  - Check-in desk
  - SuperPubs database
- Increase communication, information sharing
  - Within TechPubs
  - Within SLAC
  - Within the scientific and technical communities
- Promote use of the Internet and WWW as information resources
- Use regular staff meetings to find solutions, develop strategies collectively



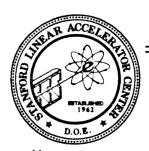
## **Impact of Budget/Staff Reductions**

- Charge-back for photos
- Conference poster service eliminated
- Rush jobs accommodated on an overtime, charge-back basis



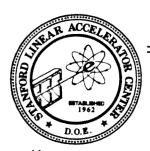
## **New Projects and Responsibilities**

- SLAC Pubs online
- Web Figures Projects
- Online Legacy Documents Project
- Taking a proactive role in the development of the SLAC Web
- TechPubs Users Group (TPUG)



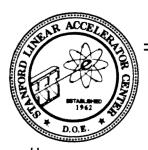
## **SLAC Pubs Online**

- Procedures now in place for concurrent publishing of all new SLAC Pubs, posting to WWW (PostScript)
- Goal: To get all SLAC Pubs on the WWW
- Exploring different online formats (Post-Script, HTML, PDF, SGML)



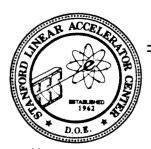
## **Legacy Documents Online Project**

- Making SLAC-internal document accessible on the WWW
  - ES&H documents
  - Administrative policies and procedures
  - Technical procedures
- Developing comprehensive database of all SLAC documents



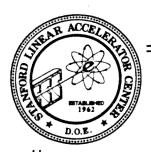
## **Web Figures Project**

- Convert all SLAC figures to portable, platform-neutral format
- Make figures publicly accessible
- Provide WWW interface to allow requestors to:
  - Preview figures
  - Select desired delivery format



## A Proactive Role in Developing the WWW at SLAC

- Some documents published on the WWW by TechPubs
  - http://www.slac.stanford.edu/grp/ techpubs/home.html
  - http://www.slac.stanford.edu/pubs/ slaconly/tip/tip.html
  - http://www.slac.stanford.edu/pubs/ slacdocs/lifeslac/lifeslac.html
  - http://www.slac.stanford.edu/pubs/ slacdocs/handbook/handbook.html
- Helping groups/departments develop their WWW pages
  - Page design, HTML assistance
  - AFS guidance
  - Web graphics



## TechPubs Users Group (TPUG)

- Monthly or occasional meetings
- Forum to increase communication between TechPubs clients and staff
  - Get feedback from users to improve responsiveness to user needs
  - Share knowledge and expertise
  - Keep users informed of changes in policy, procedure

1996 Appreciated TEAMORECENED 1910) SIVELIANT MONERAL AND COLON

S. AC ARCHIVES COLL

SERIES SUBSERIES FOLDER

## The many of the second of the **ADCoC Actions Needed**

Today: Ratify Institutional Page Committee's recommendation on main SLAC WWW graphic.

Versions of this design will appear on:

- Three core pages (Welcome, home pages)
- Secondary pages linked to the core pages

Soon:

used, it should only be the approved image. core pages) may be graphic or text. When a graphic identifier is WWWCC will very ask the ADCoC to endorse principle of at SLAC. This identifier (usually linked back to one of the SLAC requiring an "institutional identifier" on all major group pages

# SLAC Institutional (Welcome) Page Committee

## Membership:

- M. Breidenbach
- K. Cantwell
- K.Henniss, Chair
- H. Quinn

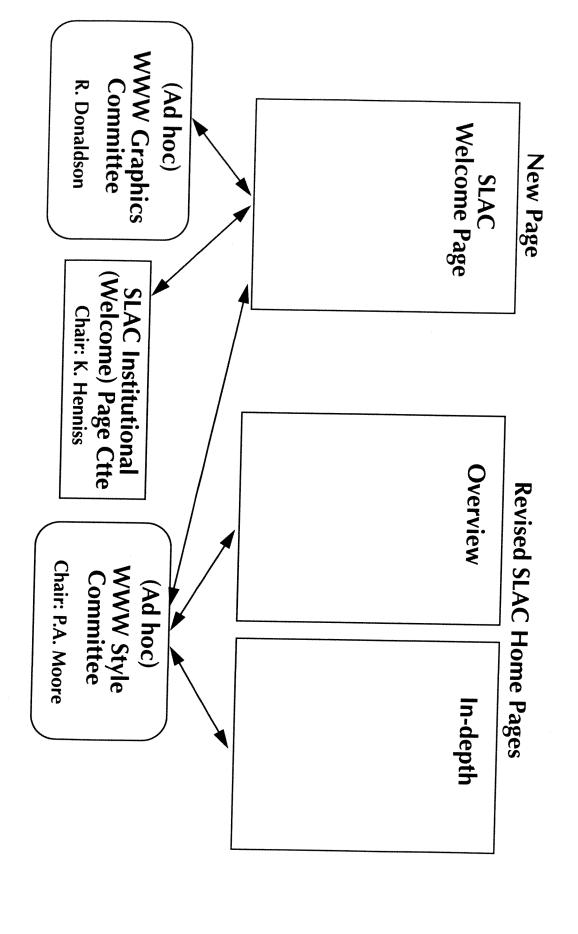
## Charge:

To assume responsibility for the pages presenting SLAC's public presence on the World Wide Web, including:

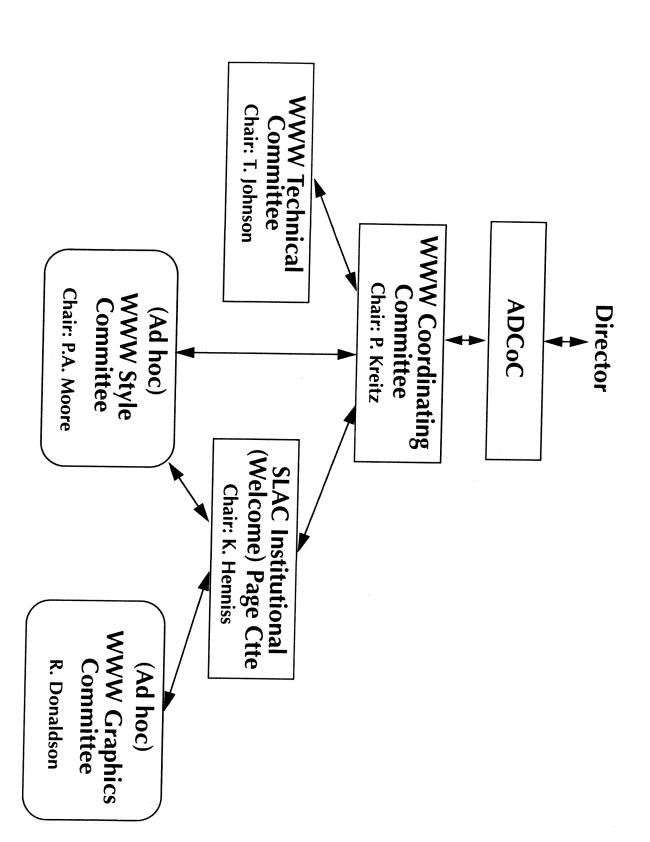
- Design (including institutional graphics)
- Content
- Coordinate efforts with VM Migration process

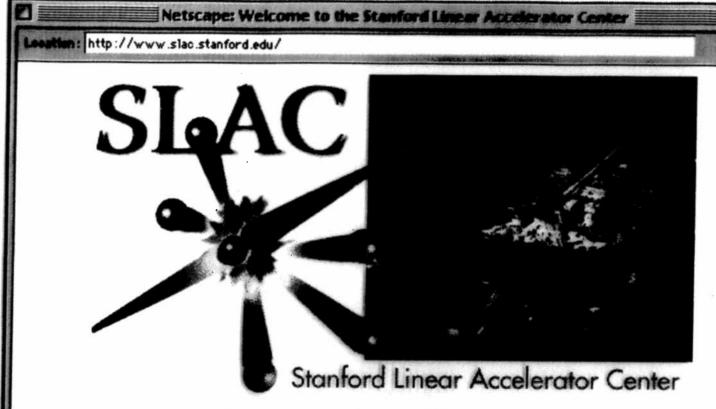
# A Structural Upgrade for SLAC's Core WWW Pages

Action: Replace SLAC's present home page with the following three-page model



# Oversight/Responsibility for the WWW at SLAC (11/95)





A national laboratory funded by the US Department of Energy.

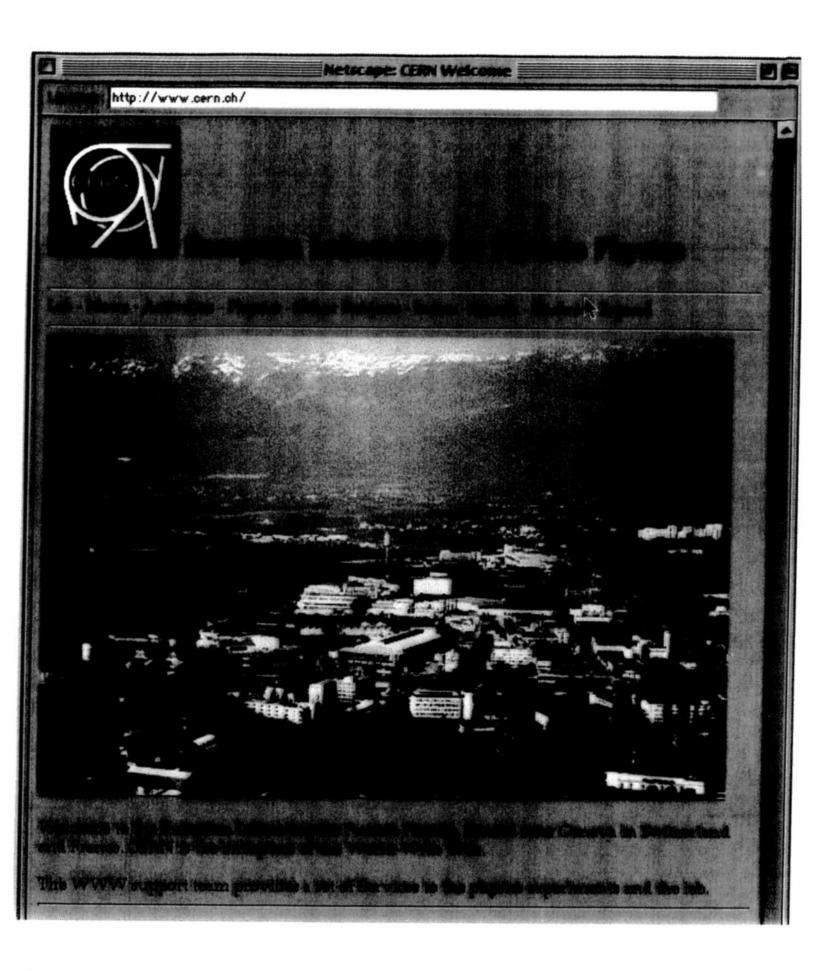
Highlighted Home Detailed Home What's New
Search Phonebook

This page provides an introduction to SLAC (home of the first U.S. WWW site) for the general public. For more details about the research programs or the organization, select either the Highlighted Home or Detailed Home button in the menubar above. Newcomers to the WWW at SLAC should see the SLAC Introduction to World Wide Web.

[Media Information]

## Welcome to SLAC

The Stanford Linear Accelerator Center (SLAC) is a national basic research laboratory, probing the structure of matter at the atomic scale with x rays and at much smaller scales with electron and positron beams. The laboratory is operated by Stanford University under a contract from the United States Department of Energy (DOE).



## TechPubs Help Pages

The UKL for this page is:

http://www.slac.stanford.edu/grp/techpubs/help/

[SLAC] [Technical Publications Department]

The following pages contain answers to some of the questions most commonly asked of the SLAC Technical Publications Department.

Contributions to the list of topics are welcome. Please send suggested topics to Kathryn Henniss. Comments/questions/suggestions about the Help pages may be sent to Maria Breaux.

## Tech Pubs Help Topics

- General TechPubs Info
  - Your Responsibilities as a SLAC Author
  - O The TechPubs Check-in Desk
  - What to Bring to the TechPubs Check-in Desk
  - Types of SLAC Documents
    - ☐ All About SLAC Pubs
    - ☐ All About SLAC Reports
- Text Processing Packages at SLAC
  - O TeX, LaTeX, and REVTeX
  - O Some Notes on LaTeX2e and LaTeX 2.09
  - Notes on Using REVTeX
  - Microsoft Word
  - o Phyzzx
- Templates
  - o For SLAC Pubs and Journals
  - O For EPAC '96 Papers
- Tips on Incorporating Figures into Your Text Documents
  - O Manipulating Figures in LaTeX2e and LaTeX 2.09 with figmacs.tex
  - o Manipulating Figures in LaTeX 2.09 with epsf.tex
  - O Incorporating, Resizing, and Placing Figures in Phyzzx
  - O Incorporating, Resizing, and Rotating Figures in MSWord 5.1
  - Incorporating, Resizing, and Rotating Figures in MSWord 6.0
  - O Creating EPS Files
- WWW Posting of SLAC Pubs, Reports, and TNs
  - O WWW Posting Information
  - O Creating PostScript Files

The TechPubs Help Pages are constantly being revised and updated. User feedback is encouraged.

Feedback to Hermiss

Last modified Thu May 30 14:03:10 PDT 1996 by Breaux.

Lecation: http://www/slac/media-info/pressphoto.html

## **SLAC Press Photos**

[SLAC]

This page makes available the most widely requested photographs of SLAC people and places. The thumbnails shown on this page are linked to high-resolution versions of the corresponding images. Clicking on a thumbnail will download the high-resolution image to your machine.

## People



Sidney Drell

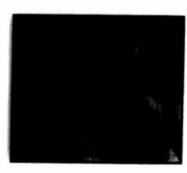
Deputy Director of SLAC and Physics Professor, Stanford University Has a distinguished career of public service in arms control and science in the national interest

Black and white image, 1996

Dimensions: 4.872 in. x 6.865 in. (1296 pixels x 1826 pixels)

Resolution: 266 pixels/inch

File Size: 2.3 MB



W. K. H. Panofsky

Director Emeritus of SLAC and Physics Professor, Stanford University An international expert on nuclear non-proliferation

Black and white image, 1996

Dimensions: 4.592 in. x 3.96 in. (1837 pixels x 1584 pixels)

Resolution: 266 pixels/inch

File Size: 2.8 MB



Martin L. Perl

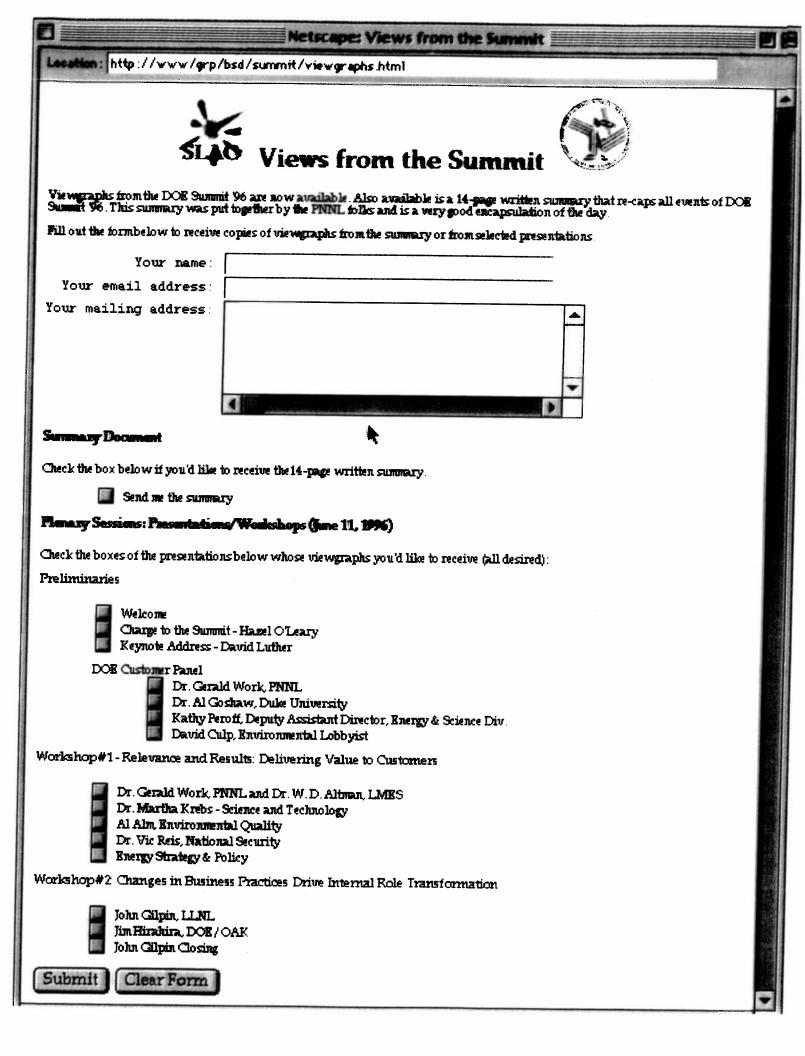
Chairman, SLAC High-Energy Physics Faculty Winner of the 1995 Nobel Prize in Physics for his 1975 discovery of the tau lepton

Black and white image, 1996

Dimensions: 4.327 in. x 6.432 in. (1151 pixels x 1711 pixels)

Resolution: 266 pixels/inch

File Size: 1.9 MB



DATE TE TO BOOK TO WISH HOW IN

SERIES SUBSERIES A FOLSER 16

## A Brief History of Posting at SLAC

## L. Trindle Gennari

Late in 1993, the Technical Publications Department's Editorial Office (SEO), at a suggestion from the WWWizards, started posting their work on SLAC's World Wide Web server. One person spent one day each week making PostScript files of the papers completed that week by the Editorial Office and posting them to the World Wide Web. There are no records of the rate of posting in the beginning, but 1 to 5 papers per week, depending on the workload in SEO, would not be far off.

In January of 1994, an additional one day of a half-time person was allotted to the posting project. This person handled all of the routine cases and shunted the problems off to the regular staff member.

In March of 1994, with 100% of the work done by SEO being posted, the procedure was expanded to include (or, at least, attempt to include) all SLAC publications, and the editors were asked to create their own PostScript files. When authors would turn in a hard copy document for printing, they were asked to provide either a PostScript file of the complete paper, including figures, or the text processor file (TeX, Microsoft Word, FrameMaker, etc.) and figure files that made up the paper. The authors were not enthusiastic about the plan to post SLAC publications, and it was a very rare individual who knew how to make a PostScript file. As a result, a very small percentage of SLAC documents were submitted in electronic form, and the source files we did receive were from all combinations of platform and text-processing package. Very quickly, the part time person was overwhelmed with the amount of work, and the number of files with unsolvable problems was growing at an alarming rate. Part of another part time person was devoted to the project to help with the load. This helped, but still didn't come close to getting the project caught up.

About this time, the SLAC Library started pulling TeX source and figures from the Los Alamos Eprint server and saving the papers to PostScript to be served on SLAC's web server. Since Los Alamos only accepted TeX that compiled without errors and only accepted figures in a given format, this was a much more straightforward procedure than the one being followed by TechPubs. Since the Library assumed that our process was as straightforward as our straightfor

Because the process used a Macintosh, source files from any platform and application could be opened and processed, but there were still many problems with fonts and included figures. Fonts not available on the processing machine would be replaced by a default font, throwing off the line lengths and formatting, usually causing disastrous page breaks and misplaced figures. Figure files created by certain older applications (usually mainframe based) would not be visible on screen (but would be visible in a printed copy), or would not appear in the correct position, or would run over the text

or otherwise fail, in any combination of works/doesn't work onscreen and works/doesn't when printed. Processing time was also very slow, because we were using a Macintosh II (!!!) to deal with the files. It was not uncommon for the reformatting and processing of one document to take an entire day.

Many of the TeX files would not process without errors (and much hitting of the continue key), and the majority of the non-TeX files would not retain their formatting when opened on our system. What, in theory, should have been 1) Open file 2)Print to PS file 3) Post, more often than not turned into an all-day ordeal of fighting with formatting, fonts, and figures before a PostScript file could be generated.

The biggest problem with accepting source files is staffing. The person processing the files must be expert in TeX, LaTeX, Microsoft Word, FrameMaker, and a host of cross-platform graphics-inclusion issues. When the volume of work required more staff, it was impossible to find a person with those skills who was willing to do a low-paying, part-time, repetitive task. The result was a string of part-time people who had to ask lots of questions before and during every step of the process.

Once a PostScript file had been created and transferred to Unix, there was a yet another set of things that could go wrong, again mostly due to included figures. The department hired a programmer to look into the PostScript problems and write some scripts for dealing with them. The most common problem was that PostScript files produced from Microsoft Word would only display any given page up to the end of the first figure on that page. Ghostview was reading the %%EOF comment at the end of the included figure file, within the PostScript file, and would interpret it as the end of the page. The fix-it script written to deal with this problem simply added an additional percent sign before the %%EOF comment. Another script would tattoo the publication number and funding acknowledgment on the first page of the file. Both scripts were dependent on specific versions of applications and printer drivers and promptly broke when new versions of the software and printer drivers were released. Since fixes like this are so fragile and require constant support to keep up with new software releases, we did not continue using this approach. As a result, many papers were posted without SLAC Pub numbers or dates.

As soon as we it was possible to do so without jeopardizing the project, authors and other paper preparers were asked to provide PostScript rather than source files. This change shortened the processing time per paper dramatically, a blessing since the volume was increasing. At this point, we were only getting files from the more advanced clients, so about half of the PostScript was usable (this dropped dramatically as more authors complied with the request). The only remaining problem was convincing authors that they really did have to turn in a PostScript file. This was solved for the most part when the DOE expressed their desire that we have an electric archive of our publications.

Now that the technology of the World Wide Web has caught up with our dream of an electric repository of SLAC publications, we are working to take this process to the next

level. In addition to continuing to collect PostScript files from authors of current publications, we have started an additional project using Adobe Acrobat Capture to scan old SLAC Pubs into PDF for delivery on the World Wide Web. Capture combines optical character recognition technology with the built-in intelligence (text is text and figures are figures) of the PDF format.

## What is PDF?

Portable Document Format (PDF) is the native format of the Adobe Acrobat development suite, a leading online document production package. PDF is closely tied in form and function to PostScript. PostScript is a programming language that is particularly suited to telling a printer where to place ink on a page. PDF does the same onscreen. In addition, PDF includes many features ideal for online delivery, including hyperlinks, automatic compression, and indexability. PDF's file compression typically makes a PDF file, which is distilled from a PostScript file (see What is Distiller?), about 80% smaller than its parent PostScript file. The compression ratio is often much higher. Table 1 shows file sizes for SLAC-PUB-7447, available from SLAC's Web site.

i Siles	
File Format	File Size
Uncompressed PostScript	12.6 Mbytes
Unix-compressed Post- Script	10.6 Mbytes
PDF	4.1 Mbytes

**Table 1: Comparison of File Sizes** 

## What is Distiller?

Distiller is a component of the Acrobat suite that parses PostScript and outputs PDF. Distiller is a very stringent PostScript parser. Irregularities passed over by more lax PostScript utilities are cause for an aborted job. Problems also arise in font inclusion and substitution. If fonts used in the PostScript file are not included in the file itself (i.e., the file expects to get the font descriptions from the printer), Distiller attempts a best guess. The results are unnattractive and unfaithful to the original at best and illegible at worst.

## What is Capture?

Acrobat Capture was developed to convert paper documents to electronic form for delivery to the Web. Although Capture boasts several output formats (MS Word, text, etc.), PDF is its native format and the only format it produces perfectly. All of Capture's most powerful features rely on PDF. Capture differs from other Optical Character Recognition (OCR) applications in that Capture preserves all of the graphics and page formatting of the original document. Most OCR programs are geared towards pages or columns of English prose. The complexity of SLAC publications, with their equations,

symbols and figures, puts them beyond the abilities of other OCR programs.

Capture addresses the combination of text and graphics by treating every scanned page as a two-layer file. One layer is a scanned image of the page saved as a graphics file, and the other is a formatted text and graphics file whose text has been processed by Capture's OCR engine. The three modes of Capture processing differ in their use of these two layers.

## **Image Only**

In Image Only mode, the text layer is discarded, and the image layer saved as a PDF file. Although the image is automatically compressed, the file is rather large, averaging 50 KB per scanned page. Since the text layer is discarded, the text is not accessible, searchable, or indexable. The result is the equivalent of a smaller, bitmapped PostScript file, accessible across a wider variety of platforms.

### **Normal PDF**

In Normal PDF mode, the text layer is placed on top of the image layer to create an Acrobat Capture document (ACD) file. The user can then use Acrobat Reviewer to confirm and edit the output of the OCR engine. Unrecognized strings are marked as suspect. The user either corrects or confirms each suspect. If the suspect is an equation, symbol, or figure, the user can cut a hole around the suspect in the text layer to allow the scanned image to show through. When the ACD file is saved to PDF, windows cut out of the text layer and the parts of the page image not showing through those windows are discarded and the layers combined into one. The resulting PDF is much smaller than the Image Only PDF, averaging 17 KB per scanned page. Also, the text and graphics on the page are available to cut and paste into other documents, search, and index.

## Image + Text

In Image + Text mode, the image layer is placed on top of the text layer and the two are combined into a PDF file without discarding any information and without user intervention. The Image + Text PDF displays like the Image Only PDF, a complete and faithful bitmapped representation of the original page. In addition, the text is accessible for indexing just as it is in Normal PDF mode. Because of this combination of layers, the resulting PDF file is larger than those for Image Only and Image + Text, averaging 53 KB per scanned page. This cost in file size is more than made up for by reduced person hours and the elimination of the human errors possible in the review process.

When TechPubs embarked on the Capture Project, we considered all three modes. We discarded Image Only out of hand because of its minimal advantages over bitmapped PostScript. Of the remaining two modes, PDF Normal appeared to hold the most potential: the maximum amount of accessible text with minimum file size.

Several weeks of trials produced the following data:

**Table 2: PDF Normal** 

Action	Tme/Size Per Page
Scan a page	20 sec
Process a page	40 sec
Review a page	15 min
Administrative overhead	5 sec
Disk space	17 KB

Using this mode, TechPubs encountered errors in the final PDF, which resulted from a combination of machine and human error. As with all OCR programs, Capture is not perfect. The errors encountered were mostly words incorrectly recognized by Capture and not highlighted as suspects for the operator to review. Capture allows the user to add words to a custom dictionary to alleviate this problem, but to eliminate it completely, the user would have to review the entire text rather than just the suspect words. As more pages are scanned, the custom dictionary grows more comprehensive and Capture becomes more accurate, however the number of person hours required to reach 100% accuracy is prohibitive.

Table 3: Ideal conditions for OCR

Font size of 10 pt or larger	
Sans serif font	
First-generation, one-sided original	
Simple English text	

TechPubs instigated a quality check to catch these errors, which brought the final PDF up to necessary quality standards. In fact, in some instances, the PDF was had few typos and misspellings than the original document. This quality check required that the PDF be printed and compared to the original. This step added a minimum of 3 minutes per page of review time and doubled the administrative overhead.:

Table 4: PDF Normal with Quality Check

Action	Tme/Size Per Page
Scan a page	20 sec
Process a page	40 sec
Review a page	18 min

Table 4: PDF Normal with Quality Check

Action	Tme/Size Per Page
Administrative overhead	10 sec
Disk space	17 KB

Reviewing and editing a PDF Normal document is essentially proof reading. The nuances of the Capture review software and the language of SLAC documents require about two weeks of full-time scanning and reviewing for an experienced proofreader to develop profficiency. The job requires constant attention to detail and is highly repetitive and often boring. The inherent problem with producing PDF Normal is that a qualified operator will not remain long in the position. (See human factors.)

## **Human Factors**

PDF Normal mode is extremely tedious work. Hours are spent staring at a screen, proofing hundreds of pages of high-energy physics scholarship (which is meaningless to the average proof reader), and listening to the loud engines of the scanner and document feeder. Couple this tedium with the job qualifications of attention to detail, intelligence, problem solving skills, familiarity with TechPubs's and Library processes and cross-platform abilities, and the result is a position that is very difficult to fill and impossible to keep filled. Also, tedium equals human error. Maintaining the level of attention to detail required for hours on end day after day is impossible.

Faced with these issues and the volume of material needing to be scanned, TechPubs began using the Image + Text mode.

Table 5: Image + Text

-
Tme/Size Per Page
20 sec
40 sec
2 sec
3 sec
53 KB

The PDF is viewed page by page to ensure that none of the pages were skipped or rotated, but no other post-scanning processing is required.

In addition to the reduction of processing time, Image + Text offers several other advantages:

 The process is extensible with occasional help (summer students). Training and startup are completed in a matter of hours.

 The process does not require constant human attention. The operator can be occupied otherwise, only checking and feeding the scanner occasionally.

• The process nearly eliminates human error. Since each page is a complete and faithful facsimile of the original, the operator needs only to confirm that each page is legible, oriented correctly, and in the correct order.

Although the per page file size is larger, disk space is less expensive than staff time, and what is lost in file size is made up in accuracy, and the size of the resulting PDF is still smaller than the same file in uncompressed PostScript.

Even more important than file size is accessibility. PDF files can be accessed by PCs, Macintoshes, and Unix machines. Since compressed PostScript is only accessible to desktop machines with commercial decompression software, some processing by the user, and a (usually unreliable) PostScript viewer, most desktop-based users at SLAC have not been able to use the compressed PostScript files on SLAC's server. To conserve network bandwidth, large PDF files have been broken up into smaller, more manageable chunks. Unfortunately, this is not possible for the corresponding large compressed PostScript files. Also, TechPubs is investigating the latest generation of PDF which provides smaller file sizes and byte-serving technology, which allows PDF documents to be sent to the browser one page at a time rather than all at once.

Potential problems with submissions: (Do we want this section?) Paper

Front pages
PostScript

Does not view
Does not print
Bad figures
Missing fonts
Does not Distill
Extra Text