

Here is a brief description of the set of requests recognized by ListProcessor
 Everything appearing in [] below is optional; everything
 appearing in <> is mandatory; all arguments are case insensitive. The vertical
 bar (|) is used as a logical OR operator between the arguments. Requests may
 be abbreviated, but you must specify at least the first three characters.
 Keep in mind that when referring to a <list>, that list may be of two kinds:
 local or remote, unless otherwise noted. When referring to a local list, your
 request will be immediately processed; when referring to a remote list (a list
 served by another ListProcessor which this system knows about), your request
 will be appropriately forwarded. Issue a 'lists' request to get a listing of a
 local and known remote lists to this ListProcessor.

ListProcessor 6.0

Below is an excerpt of the HELP file

HELP

For information on this service and how to use it, send the following
 request in the body of a mail message to listproc@info.cern.ch:

assuming that you keep the same password.

SET WWW-ANNOUNCE ADDRESS WANN new-address

You may change the address you are subscribed with
 with the following request:

SET WWW-ANNOUNCE PASSWORD WANN new-password

WARNING: Do not use your login password; you will be breaching security at you
 site.

Your initial password is WANN. You may change it
 by issuing the following request to listproc@info.cern.ch:

Any problems with this software or suggestions on the configuration
 should be mailed to me (secret@info.cern.ch)

We have now switched to ListProcessor, thanks to the suggestions
 of Richard Wiggins (six months ago :-), Lou Montulli and Mustata Akgul.

First, forget my previous message, I didn't check where the reply would go.
 I'm very sorry for that

Errors-To: secret@www0.cern.ch
 Date: Thu, 10 Feb 1994 10:55:56 +0100
 From: Arthur Secret <secret@hpww.cern.ch>
 Subject: Our new mail software
 Sender: www-announce@www0.cern.ch
 To: Multiple recipients of list <www-announce@www0.cern.ch>
 Errors-To: secret@www0.cern.ch
 Reply-To: www-announce@www0.cern.ch
 X-Envelope-To: BBBOUNNIXHUB.SLAC.STANFORD.EDU
 Content-Transfer-Encoding: 7BIT
 Originator: www-announce@info.cern.ch
 Precedence: bulk
 X-Listprocessor-Version: 6.0c -- ListProcessor by Anastasios Kotsikonas
 Content-Length: 5993

Recognized requests are:

```
help [topic]
-----
Without arguments, this file, otherwise get specific information on the
selected topic. Topics may also refer to requests. To learn more about this
system issue a 'help listproc' request. To get a listing of all available
topics, generate an error message by sending a bogus request like 'help me'.
```

```
set <list> [<option> <args>]>
```

Without the optional arguments, get a list of all current settings for the specified list. Otherwise change the option to a new value for that list. Issue a 'help set' request for more information.

```
subscribe <list> <your name>
```

The only way to subscribe to a list.

```
unsubscribe <list>
signoff <list>
```

Two ways of removing yourself from the specified list.

```
recipients <list>
review <list>
```

Get a listing of all non-concealed people subscribed to the specified list.

```
information <list>
```

Get information about the specified list.

```
statistics <list> [(subscriber email address(es)) | [-all]]
```

Get a listing of non-concealed subscribers along with the number of messages each one of them has sent to the specified list. If the optional email address are given, then statistics will be collected for these users only. For example stat foo user1@domain user2@domain will generate statistics about these two subscribers. "-all" lists statistics for all users that have posted on the list (whether currently subscribed or not).

```
run <list> [<password> <cmd> [args]>]
```

Run the specified command with the optional arguments and receive the output from stdout and/or stderr. To get a listing of all available commands to run, omit the arguments, i.e. issue a 'run <list>' request. You have to belong to the specified list, and must have obtained the password from the list's owner; the owner's address may be found in the Errors-To: header line of each delivered message. <list> may be local only.

```
lists
```

Get a list of all local mailing lists that are served by this server, as well as of all known remote lists.

index [archive | path-to-archive] [/password] [-all]
Get a list of files in the selected archive, or the master archive if no archive was specified. If an archive is private, you have to provide its password as well.

get <archive | path-to-archive> <file> [/password] [parts]
Get the requested file from the specified archive. Files are usually split in parts locally, and in such a case you will receive the file in multiple email messages -- an 'index' request tells you how many parts the file has been split into, and their sizes; if you need to obtain certain parts, specify them as optional arguments. If an archive is private, you have to provide its password as well.

view <archive | path-to-archive> [/password] [parts]
Same as "get" but in interactive mode just catenates the file on the screen.

search <archive | path-to-archive> [/password] [-all] <pattern>
Search all files of the specified archive (and all of its subarchives if -all is specified) for lines that match the pattern. The pattern can be an egrep(1)-style regular expression with support for the following additional operators: '~' (negation), '|' and '&' (logical OR and AND), '<' '>' (group regular expressions). The pattern may be enclosed in single or double quotes. Note: . matches any character including new line.

release
Get information about the current release of this listprocessor system.

which
Get a listing of local mailing lists to which you have subscribed.

Arthur

Arthur Secret,
Technical Student at CERN, phone:(41-22) 767-37-55
e-mail: secret@ddx.cern.ch

Date: Thu, 10 Feb 1994 18:35 -0800 (PST)
From: "Joan M. Winters" <WINTERS@slacvm.slac.stanford.edu>
Subject: WWW Access Issues
To: cottrell@slac.stanford.edu
Cc: wbj@unixhub.slac.stanford.edu, beb@slac.stanford.edu,
dickens@slac.stanford.edu, gbg@slac.stanford.edu,
boeh@slac.stanford.edu, ler@slac.stanford.edu,
sandra@slac.stanford.edu, leith@slac.stanford.edu,
X-Envelope-To: BEBO@UNIXHUB.SLAC.STANFORD.EDU,
BOEHM@UNIXHUB.SLAC.STANFORD.EDU, cottrell@UNIXHUB.SLAC.STANFORD.EDU,
sandra@UNIXHUB.SLAC.STANFORD.EDU
Content-Transfer-Encoding: 7BIT

Thanks for your 10 Feb 1994 16:46 update on the managers' meeting re WWW
privacy and your subsequent interchange with Leith.

I agree that 1) information resources from many parts of SLAC are
likely to be appropriate to add to the SLAC portion of the
web and 2) issues of differing access requirements by SLAC users,
collaborators, and tourists are increasingly pressing.

I have already tried to address a number of these problems working
with information owners and the WWWizards Committee, but I agree that
the time has come that we need to look at these issues on a
Lab-wide (and to some extent, HEP-wide and Web-wide) basis.

I agree that we need to investigate the resources required to support
SLAC use of WWW and how to manage them. I think we should also think
about the various SLAC needs Web services can reasonably address.

I would very much like to be involved in WWW policy efforts.

Re some specifics of your note, I would like to hear more details
about whatever problems the senior physicist encountered.

Re disk space needed by WWW, it's an interesting question since much
of the data viewable via WWW are not WWW-specific. They are also
accessed in other ways, e.g., from various data bases or
usr/local/doc files. Other files, especially images, etc., are
primarily accessed only through the Web.

Re privacy, I think we should emphasize evaluation of the NCSA server
since that purports to provide much better access control than our
current CERN server. If the NCSA server does come out as an improvement,
then we should migrate to it as soon as possible.

I will look for you tomorrow to discuss these issues further.

Joan

Date: Thu, 10 Feb 1994 16:46:42 -0800

From: COTTRILL@slac.stanford.edu
Subject: Re: Palo Alto City Links
To: winters@slac.stanford.edu
Cc: wbj@slac.stanford.edu, beb@slac.stanford.edu, dicken@slac.stanford.edu, wbj@slac.stanford.edu, gbg@slac.stanford.edu, boehlm@slac.stanford.edu, teri@slac.stanford.edu, sandra@slac.stanford.edu, leith@slac.stanford.edu
Reply-To: cottrell@slac.stanford.edu
X-Envelope-To: BEBO@UNNIXHUB.SLAC.STANFORD.EDU,
BOEHLM@UNNIXHUB.SLAC.STANFORD.EDU, sandra@UNNIXHUB.SLAC.STANFORD.EDU,
WBJ@UNNIXHUB.SLAC.STANFORD.EDU
Content-Transfer-Encoding: 7BIT

You asked about the WWW privacy issue that was on the agenda at the SCS managers meeting on Monday.

It was discussed at the managers meeting.
It was agreed it is a lab wide issue involving people like the library, public affairs, as well as any owners of data that is visible via WWW. One thought that was raised was to make a presentation to the key managers to raise the issue.

I talked to David Leith this afternoon about it. He was aware of the issue and related an example of something that happened to one senior physicist recently. He would like us to think about how to address the concerns we have raised, and respond to him. He suggested that it might be appropriate to set up a WWW user group at SLAC to address WWW issues of resources needed (e.g. disk space, cpu cycles etc.), privacy, how to manage etc. This would be more of a policy group than the more technical implementers WWW Wizards group.

In the meantime I think it would be premature to get SLAC added to the Palo Alto Resources page.

~~~~~ Forwarded from winters@slac.stanford.edu ~~~~~  
Any word from the managers' meeting re the link to the SLAC Home Page from the Palo Alto City ~other resources page?

--  
Les Cottrell

Received: from SIACVM.BITNET by SIACVM.SIAC.STANFORD.EDU (Mailier R2.08 R208004)  
with BSMTP id 3760; Sun, 06 Feb 94 16:41:45 PST  
Date: Sun, 6 Feb 1994 15:56:00 -0800  
Reply-To: "H. Galic" <GALIC@SIACVM.BITNET>  
Sender: "SIAC worldwideweb (WWW) Discussion List" <WWW-L@SIACVM.BITNET>  
From: "H. Galic" <GALIC@SIACVM.BITNET>  
Subject: Re: New SIAC Front Page  
X-To: WWW-L@SIACVM.SIAC.Stanford.EDU  
To: Bebo White <BBEO@SIACVM.BITNET>  
In-Reply-To: WINTERS@SIACVM.SIAC.Stanford.EDU -- 02/04/94 19:30

It's great that Joan is taking constant care of our Home Page.  
Even when I (or better, my laziness) protest against her rather  
firm rules on updating the HISTORY and COMMENTS pages/lines related to the  
Home Page, deeply inside I feel that she is right and that the changes must  
be well documented.

I looked with interest to the proposed  
latest version of the SIAC Home Page. It made me thinking of something  
you have perhaps discussed earlier (at the time I was not involved in this  
project): I felt very uncomfortable with the presence of the items  
"Palo Alto Weekly", and "Palo Alto City" on our Home Page. They were placed  
even higher on the page than CERN, DESY, and other physics information.

I thought we were paid by DOE to do and support PHYSICS and closely RELATED  
subjects, not to explore and advertise WWW's funny/interesting/amazing/  
other-than-physics-related features. Items which have nothing to do with  
physics (+related subjects, including computing) should not -- in my  
opinion -- appear on our Home Page. There is a place for funny and other  
interesting stuff: "Grab-bag", and "Hot-list", so let us have  
Exploratorium, and Palo Alto items there.

If I am "Johnny-came-late" in discussing this topic, and you have  
already made decisions (on some of early meetings)  
then just simply ignore my objection. If not, can we open a discussion  
(perhaps on the next WWW meeting) on what fits and what not onto slac.html.

://harv.galic

Date: Mon, 07 Feb 1994 09:57 -0800 (PST)  
From: KONRAD@ucbcsmsa.bitnet  
Subject: WWW Servers on VM  
To: BEBO@slac.stanford.edu  
X-Envelope-To: BEBO@UNIXHUB.SLAC.STANFORD.EDU  
Content-Type: TEXT/PLAIN  
Content-Transfer-Encoding: QUOTED-PRINTABLE  
Comments: Resending previously received note.

Bebo,  
Thank \*you\* for your note and, yes, I too am late in responding.  
> Sorry that it's taken a while for me to respond.

> My advice to you re: WWW on VM is don't do it.

I can understand this, having been in the C/360 and X windows development area a couple years ago, as you probably recall. Indeed, in checking around, there doesn't seem to be anything on the horizon other than unix-like stuff. NCSA has some scripts, like a REXX script, but I'm not sure how full featured it is.

Nevertheless, I still feel (and newspaper articles about America Online, CALPOLY, and others are beginning to bear out) that when network-based information services really start to accommodate large numbers of incoming clients, it won't be 20 or 30 Suns tied together that will be able to do the job - it will be an industrial strength I-wont-use-the-term-mainframe "megaserver" that will be required.

It's pretty certain that we are getting into this scale of stuff, VM or not. But I definitely would like to take you up on your offer:  
> If you pursue installing a VM server, I would welcome the  
> opportunity to collaborate with you. I'm sure that I could  
> benefit from your expertise.

More to come .....

Alan,  
Thankx,  
Allan  
konrad@ucbcsmsa.berkeley.edu (LBL)

=====  
=====

> Date: Mon, 31 Jan 1994 15:42:48 -0800  
> From: BEBO@SLAC.Stanford.EDU  
> Subject: WWW Servers on VM  
> To: konrad@ucbcsmsa.bitnet  
> Message-id: <9401312342.AA42719@hebe.SLAC.Stanford.EDU>  
> X-Envelope-to: konrad@ucbcsmsa.bitnet  
> Content-transfer-encoding: 7BIT  
> Allan,  
> Sorry that it's taken a while for me to respond.  
> SPICEL was the first WWW server on VM; it's only used now  
> for testing, so probably no one off site should point to it.  
> WWW and WWWTEST are the two primary WWW servers.  
> Both of these machines are running hacked-up versions of the

> CERN server code (not quite 2.12). The code is HTTP 1.0  
> compliant (i.e., has no problems talking most HTTP 1.0  
> browsers) but is really HTTP 0.9.  
> I used C/370 in building the server; it's my understanding  
> that CERN used Waterloo C.  
> The same is really true for the CERNVM server. The only reason we e=  
VM  
> server is for SPIRES support. When SPIRES  
> becomes available on Unix, my hope is that this server will  
> be retired. I have no plans to upgrade the VM server giving  
> this future direction.  
> My advice to you re: WWW on VM is don't do it. The original  
> CERN server code had DEFS for VM, but the latest doesn't.  
> That's why there are so many hacks. Trying to compile the  
> NCSA code is something that I haven't had the courage to  
> attempt.  
> If you pursue installing a VM server, I would welcome the  
> opportunity to collaborate with you. I'm sure that I could  
> benefit from your expertise.  
> --  
> Bebo White  
> Stanford Linear Accelerator Center  
> e-mail: bebo@slac.stanford.edu  
> phone: 415-926-2907  
> fax: 415-926-3329  
> =FF=FF WWW Servers on VM

Errors-To: secret@www0.cern.ch  
 Date: Fri, 11 Feb 1994 08:34:07 +0100  
 From: nholz@ciweng.carleton.ca (Neal Holtz)  
 Subject: Identifying scripts by file extension?  
 Sender: www-talk@www0.cern.ch  
 To: Multiple recipients of list <www-talk@www0.cern.ch>  
 Errors-To: secret@www0.cern.ch  
 Reply-To: nholz@ciweng.carleton.ca  
 X-Envelope-To: BEBO@UNNIXHUB.SLAC.STANFORD.EDU  
 Content-Transfer-Encoding: 7BIT  
 Originator: www-talk@info.cern.ch  
 Precedence: bulk  
 X-Listprocessor-Version: 6.0c -- ListProcessor by Anastasio Kotsikonas  
 Content-Length: 3283

I think this was beat to death a month or two ago, but ...

Currently, the only accepted way to designate a script gateway in a URL is to use some magic prefix such as "/cgi-bin/myscript/" to the path. Unfortunately, this co-exists very poorly with the rules for forming full URL's from relative ones (see below). What is the general opinion on using the file extension to identify gateways or scripts?

I have a small patch to NCSA HTTPD 1.1 which allows srm.conf entries like:

ScriptExtension .xyz ZZ

Then, path names that end in ".xyz" will be handled as though they had "ZZ" prepended, provided:

- a) the path hasn't been aliased, and
- b) the path doesn't already start with a script alias, and
- c) The path, after prepending "ZZ", starts with a script alias.

Therefore, with an entry like

ScriptExtension .gml /htbin/tutorial

a path like

/~nholz/91.111/q1.gml

would be handled as though it were

/htbin/tutorial/~nholz/91.111/q1.gml

Anybody else think this is desirable?

BACKGROUND (long) (for those who aren't convinced)

I am starting to develop a large document (set of lecture notes + study guide) that has quite a mixture of node types. Many of the HTML files require some filtering before being delivered to the client: one gateway filter handles delivery of tutorial style questions and the responses to them; another is sort of an "outline processor" that reveals more of the text, in line, when certain links are selected.

The basic problem is in the relative URL's that I would like to use. Having the script identifier at the front of the path name means that I have to use absolute URL's a lot, and that is going to be a real pain in a few weeks when I want to move things around. Identifying

the script in the last component of a path means that I can use relative URL's everywhere, which makes life a lot easier.

That script identifier is a type specification of a node (actually, it is both a type specification and instructions on how to convert the type of the node into one of the standard types). It has been decided, long ago, that type information will not be carried by the underlying file system. It also seems to be undesirable to have the file contents themselves reveal the type information. Therefore that type specification has to be part of the name (URL).

If the type information is encoded in the front portions of path names, and if the rules for generating full URL's from relative ones call for replacement of the rear portions of the path name, then it is impossible to use a relative URL to link to a node of a different type (because the replacement rules wouldn't allow replacement of the type information).

But of course, the rear portion (extension) is already used to convey some type information (".gif" for certain types of images, etc). It is only gateway scripts that are currently forbidden to use this mechanism.

What I am proposing is simply to allow the identification of scripts to be constant with the identification of other types of data - to be able to use the file extension as that mechanism.

Prof. Neal Holtz, Dept. of Civil Eng., Carleton University, Ottawa, Canada  
Internet: nholtz@civeng.carleton.ca Ph: (613)788-2600x5797 Fax: (613)788-3951

Received: from UCSFVM.UCSF.EDU by SLACVM.SLAC.STANFORD.EDU (Mailier R2.08 R208004) with BSMTP id 0598; Fri, 11 Feb 94 14:57:49 PST  
Received: from UCSFVM by UCSFVM.UCSF.EDU (Mailier R2.08 PTF008) with BSMTP id 7099; Fri, 11 Feb 94 14:58:13 PST  
Received: from library.ucsf.edu by UCSFVM.UCSF.EDU (IBM VM SMTP V2R1) with TCP; Fri, 11 Feb 94 14:58:13 PST  
Received: from ora.com (ruby.ora.com) by library.ucsf.edu with SMTP id AA27004 (5.67a8/IDA-1.5 for <sig-web@library.ucsf.edu>); Fri, 11 Feb 1994 14:32:29 -0800

Received: from rock.west.ora.com by ruby (8.6.5/Received: by rock.west.ora.com (5.65c/Spike-2.1) id AA17616; Fri, 11 Feb 1994 14:34:43 -0800  
Message-Id: <199402112234.AA17616@rock.west.ora.com>  
To: sig-web@library.ucsf.edu  
Subject: Volunteers and getting organized  
Date: Fri, 11 Feb 94 14:34:42 -0800  
From: "Christopher J. McRae" <mcrae@ora.com>

Thanks to all who showed up at the meeting this last time. We had about 122 people there, which is up from 80 at our previous meetings! Thanks especially to those who have volunteered to help with SIGWEB and/or Richmond High School.  
I would like to set up a special meeting with the SIGWEB volunteers to brainstorm on what-to-do and decide who's going to do it. Next week, if possible - say Wednesday or Thursday afternoon/evening for 2-3 hours? We need a place to do it, any suggestions?  
I think the first priority is getting two teams together - one will be responsible for producing the monthly meetings, and the other will be responsible for developing the sigweb online information service.  
Is there anyone out there who would like to coordinate the volunteers? This would entail getting a sigweb-volunteers mailing list going, setting meeting dates and finding a place, matching up people with projects, documenting the various efforts, etc. Of course, you don't necessarily have to do it all yourself, as long as you know how to delegate :).  
It would probably be best if you sent replies directly to me rather than the entire list, unless what you have to say is of general interest of course.

Chris

Christopher McRae  
President, SIGWEB  
Project Manager  
mcrae@ora.com  
O'Reilly & Associates, Publishers  
510/540-6036

415/242-9623

Received: by SLACVM (Mailier R2.08 R208004) id 5097;

Mon, 14 Feb 94 13:39:03 PST

Date: Mon, 14 Feb 1994 13:28 -0800 (PST)

From: "Louise Addis"

To: ADDIS@SLACVM, BEBO@SLACVM, TONYJ@SLACVM, CRANE@SLACVM, WINTERS@SLACVM,  
GALIC@SLACVM, CLANCEY@SLACVM, PKREITZ@SLACVM, JXH@SLACVM, COTTRILL@SLACVM  
Subject: problem with WWW access to SLACVM

All attempts to access port 80 on slacvm failed this morning. George  
Crane traced the problem to a full 191 disk on account WWW. He archived  
the humungous log file and erased it...solving the problem temporarily.

Patrick, could you please arrange to give WWW a much larger 191 disk...

Let's make it at least 30 cylinders. Then could you please work with Bebo  
to arrange for automatic analysis and archiving of the log file at appropriate  
intervals? TonyJ tells me that the analysis program which the guy from HCC did  
is excellent.

The hugh increase that we're seeing in HEP access via WWW (doubled from  
August to over 40,000 accesses in January...) is probably making the current  
log erasure schedule inadequate.

Received: from SLACVM by SLACVM.SLAC.STANFORD.EDU (Mailier R2.08 R208004) with BSMTP id 8634; Mon, 14 Feb 94 19:10:14 PST  
 Received: from unixhub.SLAC.Stanford.EDU by SLACVM.SLAC.STANFORD.EDU (IBM VM SMTP V2R1) with TCP; Mon, 14 Feb 94 19:10:12 PST  
 Received: by unixhub.SLAC.Stanford.EDU (4.1/SLAC 920508)  
 from charon.SLAC.Stanford.EDU  
 id AA18108; Mon, 14 Feb 94 19:11:07 PST  
 From: jxh@unixhub.slac.stanford.edu (John Halperin)  
 Message-Id: <9402150311.AA18108@unixhub.SLAC.Stanford.EDU>  
 Subject: Re: problem with WW access to SLACVM  
 To: ADDIS@slacvm.slac.stanford.edu, clancey@slacvm.slac.stanford.edu  
 Date: Mon, 14 Feb 94 19:11:07 PST  
 Cc: bebo@slacvm.slac.stanford.edu, tonyj@slacvm.slac.stanford.edu,  
 crane@slacvm.slac.stanford.edu, winters@slacvm.slac.stanford.edu,  
 galle@slacvm.slac.stanford.edu, pkreitz@slacvm.slac.stanford.edu,  
 cottrell@slacvm.slac.stanford.edu, jxh@slacvm.slac.stanford.edu  
 In-Reply-To: <01H8VV4KFM82000602@SERV02.SLAC.STANFORD.EDU>; from  
 "ADDIS@slacvm.slac.stanford.edu" at Feb 14, 94 6:01 pm  
 X-Mailer: ELM [version 2.3 PL11]

> Subject: problem with WW access to SLACVM

> All attempts to access port 80 on slacvm failed this morning. George  
 > Crane traced the problem to a full 191 disk on account WW. He archived  
 > the humungous log file and erased it...solving the problem temporarily.  
 > Patrick, could you please arrange to give WW a much larger 191 disk...  
 > let's make it at least 30 cylinders. Then could you please work with Bebo  
 > to arrange for automatic analysis and archiving of the log file at appropriate  
 > intervals? Tonyj tells me that the analysis program which the guy from HCC  
 > is excellent.  
 > The hugh increase that we're seeing in HEP access via WW (doubled from  
 > August to over 40,000 accesses in January...) is probably making the current  
 > log erasure schedule inadequate.

Louise, Patrick -- I don't think the above is a correct diagnosis of the  
 problem, or that WW's disk space has anything to do with it. A full 191  
 disk is \*normal\* for the VM WW server, and it doesn't seem to have any  
 effect other than stopping its overly verbose logging. The problem this  
 morning seems to be just another instance of the same VM server problems  
 we've been having for quite some time now, as documented in Protrak  
 #11154. There have been 12 such crashes in the last two weeks. Please  
 hold off on expanding WW's disk space until we can discuss the problem,  
 perhaps at this week's WW wiz meeting.

John Halperin <jxh@SLAC.Stanford.Edu>

Stanford Linear Accelerator Center.

John Halperin <jxh@SLAC.Stanford.Edu>

Networking Group

Received: by SLACVM (Mailier R2.08 R208004) id 8892;  
Mon, 14 Feb 94 20:05:00 PST

Date: Mon, 14 Feb 1994 19:59 -0800 (PST)

From: CRANE@SLACVM

To: jxh@unixhub.slac.stanford.edu, ADDIS@slacvm.slac.stanford.edu,

clancey@slacvm.slac.stanford.edu

cc: bebog@slacvm.slac.stanford.edu, tonyj@slacvm.slac.stanford.edu,

crane@slacvm.slac.stanford.edu, winters@slacvm.slac.stanford.edu,

galic@slacvm.slac.stanford.edu, pkreitz@slacvm.slac.stanford.edu,

cottrell@slacvm.slac.stanford.edu, jxh@slacvm.slac.stanford.edu

Subject: Re: problem with WW access to SLACVM

In-Reply-To: jxh@unixhub.slac.stanford.edu -- 02/14/94 19:10

John,

Louise and I restarted WW and found it to still be unresponsive. The 191 disk was 99% full, primarily because of the logging file. When we erased the logging file and re-started the server everything was happy again. It should be easy enough to fill the disk up a reproduce the problem again if you like.

George

Errors-To: postmaster@ww0.cern.ch  
Date: Tue, 15 Feb 1994 11:25:40 +0100

From: Jon.Tetzchner@nta.no  
Subject: fm2html version 0.8.4 now available

Sender: ww-talk@ww0.cern.ch

To: Multiple recipients of list <ww-talk@ww0.cern.ch>

Errors-To: postmaster@ww0.cern.ch

Reply-To: Jon.Tetzchner@nta.no

X-Envelope-To: BEBO@UNIXHUB.SIAC.STANFORD.EDU

Content-Transfer-Encoding: 7BIT

Originator: ww-talk@info.cern.ch

Precedence: bulk

X-ListProcessor-Version: 6.0c -- ListProcessor by Anastasios Kotsikonas

Content-Length: 9404

Fm2html version 0.8.4 is now available. Get it from bang.nta.no. It is available in the pub directory.

Below is some information about the filter and a list of changes from the last version.

Enjoy!

Jon.

-----  
Jon Stephenson von Tetzchner  
Norwegian Telecom Research  
jons@hal.nta.no  
-----

## Frame 2 Html filter

=====  
This filter has been made as part of my work at Norwegian Telecom Research. All rights to the filter belong to Norwegian Telecom Research. The filter can be modified as long as the top lines of each source file indicating the original source of the filter is kept. We would also appreciate being sent information about bugs, bug fixes, etc. The code can not be included in a commercial product.

The current version of the filter:

- o Handles frame files and books.
- o Is customizable, through a tags file mapping frame tags to logical tags used by the filter.
- o All frame X-refs become html links.
- o An index is automatically generated based on chapter headings in the frame documents.
- o The file structure of the frame document is kept in the html document. Single frame files become single html files. Frame books become multiple html files, one html file for each frame file. FrameMaker generated files are removed.
- o Graphics and maths are separated to files, which are then translated to postscript and ultimately gif.
- o Tables are handled through the <pre> html tag.
- o Italics and bold parts of paragraphs are handled.

Problems and bugs:

- o Characters in frame not in html, including Greek character.

o Documents that are divided into sections using the Framemaker

- o Framemaker documents which do not use tags and operate on text directly.
  - o Very large pictures are not handled well (bigger than an A4).
  - o Only anchored frames are handled well. Other graphics may be lost.
- This program was made for internal use only. I will try to make it portable, but it is not unlikely that a few paths have to be changed.

#### Acknowledgement:

I would like to thank Geir Ivarsoey of Norwegian Telecom Research, Duncan Fraser and Randy Roesler of Macdonald Dettwiler, Daniel K. Schne TECRA (Educational Technologies and Learning), Jan van der Steen of Computer Lab, Amsterdam, Stephen Martin of LSI Logic Corporation of Canada, and Rob Kooper and Bob Jackson for their assistance in improving this program.

#### Availability:

ftp.bang.nta.no:pub/fm2html.v.0.n.m.tar.Z.

#### Disclaimer:

Use this program at your own risk.

Below is a list of environment variables that can be set:

FRAME2HTMLPRINTFILE (e.g. /local/lib/mf2giff.printsettings)  
FRAME2HTMLTAGSFILE (e.g. /local/lib/mf2html.tags)  
DOCUMENTLANGUAGE (US OR NO, US is default. This results in "Inhold" b used instead of "Table of Contents", etc.)

#### Usage:

fm2html frame-file [title] [author]  
The first specifies the name of the file or book to translate. The second optional parameter is the title of the file. The third optional parameter is the author of the file. fm2html can also be called instead of fm2html if you have transformed to mif already. mif2html takes the same parameters as fm2html (except the first that must be a mif file). fm2mif can be used to translate frame files to mif. It takes as its only parameter the name of the frame file.

#### Warning:

The filter operates in the current version on your local directory. It uses a lot of memory and will very easily run out of space, unless you give it a lot of swap space.

#### Some comments:

The tags files describe the connection between Framemaker tags and the internal tags of the filter. These internal

tags are mapped to HTML tags. When the filter program is run, it checks whether there exists a file called <filename>.tags, where <filename> is the name of the file or book to translate. If no such file is found, the filter checks whether the FRAME2HTMLTAGSFILE environment variable is set. If it is, the corresponding file is used. If the environment variable is not set, then the default file is used. The default file is \$LIBDIR/mit2html.tags.

The tags file uses TABs to separate internal tags and FrameMaker tags. To see a list of the internal tags, have a look at the tags file.

You may have to change the line in the start of the perl scripts identifying the location of the perl interpreter. Do this in the Makefile.

You must edit the destination location for files (LIBDIR and BINDIR) to a suitable directory. Then run 'make install'.

You will probably have to edit the directories in ps2gif so that they show where the pnm and ppm filters lie. Other programs that have to be available on your machine include gs (ghostscript), Framemaker (lmbatch), flex, yacc and cc.

You can include links to other html documents in your Framemaker document by using Markers of the type 'Hypertext'. The marker text is assumed to be an URL, e.g. "http://www.nta.no/". Any text following the marker (until the end of the line or until another marker or such in the text) becomes a link in your HTML file to the file given by the marker text. Tip: To avoid having too much text highlighted, have a marker (not Hypertext) denoting the end of the text to be highlighted. Framemaker newlink, opening and gotoLink are now supported.

Any paragraphs of the type 'TITLE' will become the title of the document and will be moved to the top of the document (or of the book document, if the file being translated is a book).

Have a look at the Demo.doc file (Demo.html) to see an example file.

Frequently asked questions:

o Where can I find the pbmpus (pnm and ppm filters) packages?

See a list of locations given by archie in "WhereToFindPbmplus

o Where can I find flex?

See a list of locations given by archie in "WhereToFindFlex"

o What is the fl (-lfl) library?

It is the flex library.

version 0.8.4

o Unlisted headings (UNHEADING) added. Do not show in the table of contents.

- o Table of contents cleaned up and changed in appearance (bold major headings and less space between lines).
- o Table Headings and title are bold.
- o Figure text is now bold.
- o Fixed links to tables bug.
- o Bold and italics are now handled in tables (works only with Mosaic 2.2 and n
- o Figures and mathematic formulas are no longer totally removed from tables. I
- a link with the tekst "(fig)" is generated. Clicking the link shows the miss

- o FrameMaker hypertext handling extended.
- o FrameMaker hypertext links (opentlink and gotoLink) are now handled (Stephen
- o Removed any tags from <TITLE> field as WinMosaic displays them.
- o Changed mfg2gt so that the files are opened and closed in order instead
- of opening all at the same time. Saves space...
- o Introduced use of TMPDIR variable instead of /tmp. /tmp is used when the
- o TMPDIR variable is not set (Bob Jackson).
- o Changed extractmthhead.pl and trans.l to follow FrameMaker 4 standard.
- (Bob Jackson).
- o Modified YERROR and changed YMAXDEPTH variable to allow the converter to
- work on SGI systems running IRIX-4.0.5 (Jan van der Steen).
- o Fixed unfinished indenting bug.
- o Norwegian Document Language now works.
- o Fixed multiple .html's in references between files.
- o New tags, some old tags dropped.
- o Updated and documented tags file.
- o Changed insertcharacters.l so that special characters are handled correctly
- when compiled with cc as well as gcc.
- o Made links inside skip filename, so that the file is not re-read.
- (Dropped because this still messes up back and forward in Mosaic).
- o Made the use of the FRAME2HTMLPRINTERFILE env variable optional
- (Duncan Fraser).
- o Prepared special characters somewhat for the acceptance of the proposed HTML
- character set (alpha; etc.).

Version 0.8.3 (Jan. 3. 1994)

- o Fixed bug with links between files in a book (Daniel K. Schneider).
- o Fixed continuing symbols bug.
- o Fixed symbols and tags bug.
- o Bugfixes. Hopefully the end of the 'Error in parsing. Got <Char whatever>.'

Version 0.8.2 (Nov. 30. 1993)

- o Fixed figure or table at end of file disappearance bug.
- o Fixed footnote in table bug.
- o Figures are now also links to themselves. This enables the reader
- to resize the figures if the viewer allows this (e.g. xv).
- o Colour figures.
- o .doc stripped (Duncan Fraser).
- o Hypertext markers in Frame Documents are assumed to be HTML links.
- The following text becomes a link to the document given by the
- FrameMaker marker. Markers can be put in tables as well. However,
- figures can not be links... except to themselves... for now anyway.
- o More bugfixes

Version 0.8.1. (Nov. 24. 1993)

- o Various bug fixes.
- o More translation.
- o Improved list handling (Duncan Fraser).
- o Fixed Makefile bugs (missing addtext).
- o Fixed bug with files ending in ".doc".

Version 0.8 (Nov. 8. 1993)

Errors-To: postmaster@ww0.cern.ch  
Date: Tue, 15 Feb 1994 12:20:00 +0100  
From: montuili@stat1.cc.ukans.edu (Lou Montuili)  
Subject: Changes to news:\* handling  
Sender: www-talk@ww0.cern.ch  
To: Multiple recipients of list <www-talk@ww0.cern.ch>  
Errors-To: postmaster@ww0.cern.ch  
Reply-To: montuili@stat1.cc.ukans.edu  
X-Envelope-To: BEBO@UNNIXHUB.SLAC.STANFORD.EDU  
Content-Transfer-Encoding: 7BIT  
Originator: www-talk@info.cern.ch  
Precedence: bulk  
X-Listprocessor-Version: 6.0c -- listprocessor by Anastasio Kotsikonas  
Content-Length: 4321

Our local news guru (Bob Sloan) pointed out to me that the Web use of "LIST" to retrieve all the articles on the system was rather ugly and that what we should be using is "LIST NEWSGROUPS". "LIST" returns the name of each group plus the last article number and the first article number and then a moderated code. "LIST" looks like:

#### NEWSGROUPS

```
* alt.3d 5025 4760 Y
* alt.abortion.inequity 12827 12537 Y
* alt.abuse-recovery 2 3 m
* alt.activism 60722 60194 Y
```

Pretty darn ugly. "LIST NEWSGROUPS" returns the name of each newsgroup and a description. With some slight code changes we now get:

#### NEWSGROUPS

```
alt.3d
Three-dimensional imaging.
alt.abortion.inequity
Paternal obligations of failing to abort unwanted child.
alt.abuse-recovery
Companion to alt.sexual.abuse.recovery. (Moderated)
alt.activism
Activities for activists.
```

I have attached code diffs for both the Lynx version of the WWW library changes and the Xmosaic 2.2 version, since there are differences in the file. I will have these changes in Lynx version 2.2

Does anyone know of any server that doesn't support "LIST NEWSGROUPS"? I tried INN, ANU News and a CNEWS installation and they all worked.

:lou

```
----- Diffs to Lynx version of libwww 2.1 HTNews.c -----
602c602
< (*targetclass.start_element)(target, HTML_DL , 0, 0);
---
> (*targetclass.start_element)(target, HTML_MENU , 0, 0);
612c612
< (*targetclass.start_element)(target, HTML_DT , 0, 0);
---
> (*targetclass.start_element)(target, HTML_LI , 0, 0);
624,637c624,630
```

```
> int i=0;  
<> /* find whitespace if it exits */  
<> for( ; line[i] != '\0' && WHITE[line[i]]; i++)  
<> ; // null body */  
<> if(line[i] == '\0') {  
<>     line[i] = '\\';  
<>     write_anchor(line, line);  
<>     (*targetClass.start_element)(target, HTML_PD , 0, 0);  
<> PUTS(kline[i+1]); /* put description */  
<> } else {  
<>     write_anchor(line, group);  
<> PUTS(line);  
-----<>  
<> char group[LINE_LENGTH];  
<> int first, last;  
<> char postable;  
<> if(sscanf(line, "%d %d %c", group, &first, &postabl  
>    else  
>        PUTS(line);
```

```

< (*targetClass.end_element)(target, HTML_DL);
--
> (*targetClass.end_element)(target, HTML_MENU);
755c748
< strcpy(command, "LIST NEWSGROUPS");
--
> strcpy(command, "LIST ");
--
*****
* THE UNIVERSITY OF KANSAS
* Lou MONTULLI @ UkanaiX.cc.ukans.edu
* Kuhnub.cc.ukans.edu
* ACS Computing Services
* 913/864-0436
* Ukanvax.bitnet
* UNIX! Cool! I know that! Jurassic Park - The Movie
*****

```

Errors-To: postmaster@ww0.cern.ch  
Date: Tue, 15 Feb 1994 12:29:54 +0100  
From: luotonen@ptsun00.cern.ch (Ari Luotonen)  
Subject: CERN httpd 2.15beta released  
Sender: ww-talk@ww0.cern.ch  
To: Multiple recipients of list <ww-talk@ww0.cern.ch>  
Errors-To: postmaster@ww0.cern.ch  
Reply-To: luotonen@ptsun00.cern.ch  
X-Envelope-To: BBBO@UNIXHUB.SLAC.STANFORD.EDU  
Content-Transfer-Encoding: 7BIT  
Originator: ww-talk@info.cern.ch  
Precedence: bulk  
X-Listprocessor-Version: 6.0c -- ListProcessor by Anastasios Kotsikonas  
Content-Length: 5730

CERN httpd 2.15beta is out:

ftp://info.cern.ch/pub/ww/src/cern-httpd\_2.15beta.tar.Z

This package includes EVERYTHING, so don't ftp the libww. Also, DO NOT use the libww that comes with it for anything else, libww 2.15 is NOT officially released.

Precompiled binaries exist for:

Sun4: ftp://info.cern.ch/pub/ww/bin/sun4/httpd\_2.15beta.Z  
Solaris: ftp://info.cern.ch/pub/ww/bin/solaris/httpd\_2.15beta.Z  
HP: ftp://info.cern.ch/pub/ww/bin/snake/httpd\_2.15beta.Z  
NEXT: ftp://info.cern.ch/pub/ww/bin/next/httpd\_2.15beta.Z  
Decstation: ftp://info.cern.ch/pub/ww/bin/decstation/httpd\_2.15beta.Z  
OSF/1: ftp://info.cern.ch/pub/ww/bin/osf1/httpd\_2.15beta.Z

For other platforms I'm happy to receive diffs. This release is mainly for CGI/1.0 and proxy gatewaying, but a lot of other new features have been included (that's why I still call it beta).

#### CERN HTTPD 2.15beta RELEASE NOTES

There is one single thing that needs to be done when changing over from httpd 2.14 to 2.15:

Rename your old /htbin scripts to end in .pp suffix!

#### General Notes

- \* Code tested under Purify -- all detected memory leaks and revealed bugs fixed.
- \* Forking code enhanced -- no longer crashes when running standalone.
- \* Documentation redesigned, but still under construction:  
http://info.cern.ch/hypertext/WWW/Daemon/User/Guide.html
- \* Contains Solaris ports! (but not VMS :- ( )

CGI/1.0, Common Gateway Interface

- \* CGI/1.0 interface fully implemented
- \* Old CERN httpd scripts will continue working if you rename them to end with .pp suffix. Links referencing these scripts do NOT need to be changed. (This feature does not add any overhead to

```

* CGI/1.0 script calls.)
* New product cgiparse for CGI/1.0 scripts to parse QUERY_STRING
  env.var and to read CONTENT_LENGTH characters from stdin
* htmimage upgraded to CGI/1.0
* The whole server-environment is propagated to CGI script, except
  for variables that are reserved for CGI/1.0.
* Scripts are spawned by doing a fork() and exec() instead of
  system() -- more efficient and secure

```

#### Firewall Gateway Modifications

```

* Access authorization works thru firewalls
* So does POST, therefore forms also
* -disable/-enable command line options and Disable/Enable
  configuration directives for dis/enabling HTTP methods. GET, HEAD
  and POST are enabled by default.
* Fix: text/html and text/plain not passed multiply to servers when
  running as gateway
* Fix: /*, image/* etc not expanded by the gateway
* Fix: try local search ONLY when accessing local files
* Known bug remaining: big binary files fail to transfer

```

#### Other New Features

```

* When started standalone in non-verbose mode automatically
  disconnects from terminal session and goes background
* User-supported directories enabling URLs starting with ~/username
  redirection
* Meta-information files to allow RFC-822-style headers to be
  appended to server response header section
* New, common logfile format, localtime default, GMT as an option
  Ability to suppress logging for certain hosts/domains according to
  given hostname template or IP number mask, like *.cern.ch or
  128.141.*.*
* -setuid option to set server uid to authenticated uid (local)
  Multilanguage support: same URL can be used to retrieve a document
  in different languages
* AddLanguage, AddEncoding and AddType directives to configuration
  file (AddType replaces Suffix -- suffix still understood)
* Better multiformat algorithm
* HostName directive to config file for servers that want to give
  CGI/1.0 scripts a different hostname than the actual. Useful if
  machine has many aliases, or if httpd fails to get the full
  domainname.
* Exec rule obsoleting HTBIn directive -- now multiple script
  directories possible, with arbitrary mappings
* Get-Mask, Post-Mask and Put-Mask for protection setup files.
  Mask-group still understood
* Groups All/Users and Anybody/Anyone/Anonymous automatically
  defined. All means anybody that has been authenticated, and
  Anybody is just anybody

```

#### Enhancements, Fixes

```

* The final explicit Map to filesystem in configuration file no
  longer required, because it was causing confusion
* Assume Basic authentication scheme even if not explicitly

```

```

* HEAD works, also with stupid scripts that also output the body
  eventually be made more general)
* Scripts can output also Uri: and Expires: headers (this will

```

mentioned in setup file  
 \* Get client DNS hostname, for the logfile among other things  
 \* Fail made the default when rules are translated to the end without  
 coming across with a Pass, Exec or Fail rule (this is to enhance  
 security, it was too easy to forget the Fail \* from the end of  
 config file)  
 \* Made config (rule) file understand different ways of writing  
 keywords, e.g.: UserDir, userdir, User-Dir, user\_dir,  
 UserDirectory and so on  
 \* The eight misplaced server-side access authorization files moved  
 away from libwww  
 \* Fix: directory indexing works with a trailing slash  
 \* Fix: HTSimplify() called strcpy() with overlapping args

Ari Luotonen  
 World-Wide Web Project  
 CERN  
 CH - 1211 Geneva 23  
 | httpdd@info.cern.ch  
 | phone: +41 22 767 8583  
 | email: luotonen@dx.cern.ch

Errors-To: listmaster@ww0.cern.ch  
 Date: Tue, 15 Feb 1994 21:28:48 +0100  
 From: "Daniel W. Connolly" <connolly@hal.com>  
 Subject: HTML DTD: past, present, future [was: HTML DTD and HYTime]  
 Sender: ww-talk@ww0.cern.ch  
 To: Multiple recipients of list <ww-talk@ww0.cern.ch>  
 Errors-To: listmaster@ww0.cern.ch  
 Reply-To: connolly@hal.com  
 X-Envelope-To: BEBO@UNIXHUB.SLAC.STANFORD.EDU  
 Content-Transfer-Encoding: 7BIT  
 Originator: ww-talk@info.cern.ch  
 Precedence: bulk  
 X-Listprocessor-Version: 6.0c -- Listprocessor by Anastasio Kotsikonas  
 Content-Length: 52426  
 -----  
 Content-Type: text/plain; charset="us-ascii"  
 Content-ID: <4639.761342909.2@lusa>  
 Content-Description: plain text version

Toward a Formalism for Communication On the We  
 Daniel W. Connolly <connolly@hal.com  
 \$Id: html\_essay.html,v 1.2 1994/02/15 20:07:12 connolly Exp connolly

Status

I had hoped to polish this more before publishing it, but I can't seem to  
 get caught up... there's so much new stuff all the time!

# SOME BACKGROUND ON SGML FOR THE WORLD-WIDE WEB

In late 1992 and early 1993, I did quite a bit of work on the HTML DTD while  
 I was working at Convex in the online documentation group.

When I began, there was the LineMode browser and the NEXT implementation,  
 and a few nodes in the web describing HTML with some oblique references to  
 SGML. I was not intimately familiar with SGML, but I was quite familiar with  
 the problems of document interchange, and I was eager to apply some of my  
 formal systems background to the problem.

## On Formally Unconvertible Document Formats

My experience with document interchange led me to classify document formats  
 using the essential distinction that some are "programmable" and some are  
 not. Most widely used source forms are programmable: Tex, troff, postscript  
 and the like. On the other hand, there are several "static" formats: plain  
 text, Microsoft RTF, Framemaker MIF, GNU's Texinfo,

The reason that this distinction is essential with respect to document  
 interchange is that extracting information from documents in "programmable"  
 document formats is equivalent to the halting problem. That is, it is  
 arbitrarily difficult and cannot be automated in a general fashion.

For example, I conjecture that it is impossible to write a program that will  
 extract the third word from a Tex document. It would be an easy task for 80  
 of the Tex documents out there -- just skip over some formatting stuff and  
 grab the third bunch of characters surrounded by whitespace. But that  
 "formatting stuff" might be a program that generates 100 words from the  
 hyphenation dictionary. So the simple lexical scan of the Tex source would  
 find a word that is not third word of the document when printed.

This may seem like an obscure and unimportant problem, but I assure you that  
 the problem of converting Tex tables to Framemaker MIF is just as  
 unsolvable.

So while "programmable" document formats have the advantage that features can be added on a per-document basis, they suffer the disadvantage that these features cannot be recovered by the machine and translated in an automated fashion.

#### Document Formats as Communications Media

If we look at document formats in light of the conventional sender/message/medium/receiver communications model, we see that document formats capture the message at various levels of "concreteness". The message begins as a collection of concepts and ideas in the mind of the sender. In order to communicate, the sender and receiver must share some language. That is, they must both understand some common set of symbols and the way those symbols combine to represent ideas. The sender's job is to express the message in terms of the common symbols and express them on the medium -- that is "render" or "present" them. The medium stimulates the receiver to reconstruct the symbols in his/her brain -- that is, the receiver "interprets" or "recognizes" the symbols from the medium. Those symbols interact with other symbols in the receiver's brain, and the receiver "gets the message."

The communications medium is often a layered combination of more and less concrete media. For example, folks first render their ideas in the symbolology of the English language, and then render those symbols as sequences of spoken phonemes or written characters. Those written characters are in turn combinations of lines, curves, strokes, and points. The receiving folks then assemble the strokes into characters, the characters into words, the words into phrases, sentences, thoughts, ideas, and so on.

The most common and ubiquitous document format, plain ASCII text, captures or digitizes messages at the level of written characters. Postscript captures the characters as lines, curves, and paths. The GIF format captures a document as an array of pixels. GIF is in many ways infinitely more expressive than plain text, which is limited to arrangements of the 96 ASCII characters.

The RTF, Tex, nroff, etc. document formats provide very sophisticated automated techniques for authors of documents to express their ideas. It seems strange at first to see that plain text is still so widely used. It would seem that Postscript is the ultimate document format, in that its expressive capabilities include essentially anything that the human eye is capable of perceiving, and yet it is device-independent.

And yet if we take a look at the task of interpreting data back into the ideas that they represent, we find that plain text is much to be preferred, since reading plain text is so much easier to automate than reading GIF files (optical character recognition) or Postscript documents (halting problem). In the end, while the source to a various Tex or troff documents may correspond closely to the structure of the ideas of the author, and while Postscript allows the author very precise control and tremendous expressive capability, all these documents ultimately capture an image of a document for presentation to the human eye. They don't capture the original information as symbols that can be processed by machine.

To put it another way, rendering ideas in Postscript is not going to help solve the problem of information overload -- it will only compound the situation.

As a real world example, suppose you had a 5000 page document in Postscript and you wanted to find a particular piece of information inside it. The author may have organized the document very well, but you'd have to print it to use those clues. If the characters aren't kerned much, you might be able to use grep or sick a WAIS indexing engine on it. Then, once you've found

what looks like postscript code for some relevant information, you'd pray that the document adheres to the Adobe Document Structuring conventions so that you could pick out the page containing the information you need and view that page.

If that's too perverse, look at the problem of navigating a large collection of technical papers coded in Tex. Many of the authors use LaTeX, and you may be able to convince the indexing engine to filter out common LaTeX formatting idioms -- or better yet, weight headings, abstracts, etc. more heavily than other sections based on the formatting idioms. While there are heuristic solutions to this problem that will work in the typical 80%/20% fashion, the general solution is once again equivalent to the halting problem; for example, individual documents might have bits of Tex programming that change the significance of words in a way that the indexing engine won't be able to understand.

#### SGML as a Layered Communications Medium

So where does SGML fit into the sender/message/medium/receiver game?

I'll use Postscript as a basis of comparison. The Postscript model consists of a fairly powerful and general purpose two dimensional imaging model, that is, a set of primitive symbols for specifying sets of points in two dimensions using handy computational techniques, and a general purpose programming model for building complex symbols out of those primitives. The model is applied extensively to the problem of typography, and there is a architecture (that is, a set of well known symbols derived from the primitives) for using and building fonts.

So to communicate message consisting of symbols from human communications to Postscript, one may choose from a well known set of typefaces, or create a new typeface using the well known font architecture, or free-hand draw some characters using postscript primitives, or draw lines, boxes, circles and such using postscript primitives, or scribble on a piece of paper, scan it, and convert the bits to use the postscript image operator. The space of symbols is nearly limitless, as long as those symbols can be expressed ultimately as pixels on a page.

The distinctive feature of Postscript (an advantage at times, and a disadvantage at others) is that whether you print it and deliver the paper or you deliver the Postscript and the receiver prints it out, the result is the same bunch of images.

The SGML model, on the other hand, specifies no general purpose programming model where complex symbols can be defined in terms of primitive symbols. The meaning of a symbol is either found in the SGML standard itself, or in some some PUBLIC document (which may or may not be machine readable), or in some SYSTEM specific manner, or defined by an SGML application. The only real primitives are the character and the "non-SGML data entity".

The model prescribes that a document consist of a declaration, a prologue, and an instance. The declaration is expressed in ASCII and specifies the character sets and syntactic symbols used by the prologue and instance. The prologue is expressed in a standard language using the syntactic symbols from the declaration, and specifies a set of entities and a grammar of element types available to the instance.

The instance is a sequence of elements, character data, and entities constrained by the grammar set forth in the prologue, and the SGML standard does not specify any semantics or meaning for the instance.

So to communicate using SGML, the sender first chooses a character set and certain processing qualities and capacities. For example "I'm writing in ASCII, and I'll never use an element name more than 40 characters long" is some information that can be expressed in the SGML declaration. [The

standard allows the SGML declaration to be implicitly agreed upon by sender and receiver, and this is generally the case].

The tricky part is the prologue, where the sender gives a grammar that constrains the structure of the document. Along with the information actually expressed in SGML in the prologue, there is usually some amount of application defined semantics attached to the element types. For example, the prologue may express in SGML that an H2 element must occur within the content of an H1 element. But the convention that text in an H1 is usually displayed larger and considered more important is application defined.

Once the prologue is determined (this usually involves considerable discussion between a collection of authors and consumers in some domain -- in the end, there may be some "parameter entities" in the prologue which allow some variation on a per-document basis), the sender is constrained to a rigorous structure for the organization of the symbols and character data of the document. On the other hand, s/he has an automated technique for verifying that s/he has not violated the structure, and hence there is some confidence that the document can be consumed and processed by machine.

#### THE HTML DTD: CONFORMING, THOUGH EXPEDIENT

Design Constraints of the HTML DTD

Tim's original conception of HTML is that it should be about as expressive as RTF. In contrast to traditional SGML applications where documents might be batch processed and complex structure is the norm, HTML documents are intended to be processed interactively. And the widespread success of MYSTIFY word processors based on fairly flat paragraph structure was proof that something like RTF was suitable for a fairly wide variety of tasks. As I learned a little about SGML, it was clear that the WWW browser implementation of HTML sorely lacked anything resembling an SGML entity manager. And there were some syntactic inconsistencies with the SGML standard. And it didn't use the ID/IDREF feature where it should have...

Then, as I began to comprehend SGML with all its warts, (who's idea was it to attach the significance of a newline character to the phase of the moon anyway?) I was less gung-ho about declaring all the HTML out there to be blasphemous to the One True SGML way.

Thus I chose for my battle to find some formal relationship between the SGML standard and the HTML that was "out there." The quest was:

FIND SOME DTD SUCH THAT THE VAST MAJORITY OF HTML DOCUMENTS ARE INSTANCES OF THAT DTD, CONVERSELY, SUCH THAT ALL ITS INSTANCES MAKE SENSE TO THE EXISTING WWW CLIENTS.

I struggled mightily with such issues as:

Should we be sticking <!DOCTYPE HTML SYSTEM> in .html files? What if somebody puts an entity declaration in there? (And does that mean that WWW clients have to be able to parse SGML prologues in general?)

What's the syntax of an attribute value? If we allow SHORTTAG YES, does that mean we have to parse <em/this/ style of markup too?

Can we put some short reference maps in the DTD that will cause real SGML parsers and current WWW browsers to do the same thing w.r.t newlines? (i.e. can we make all that phase-of-the-moon processing with newlines a moot issue)

What about marked sections? Short reference maps?

What character set should we be using? How do I express ISO-Latin-1 in

Received: by SLACVM (Mailier R2.08 R208004) Id 9768;  
Thu, 27 Jan 94 07:27:50 PST  
Date: Thu, 27 Jan 1994 07:27 -0800 (PST)  
From: "John Halperin"  
To: BEBO@SLACVM, CRANE@SLACVM  
Subject: WWW rdr files

<JXH@SLACVM>

Any idea why there are a dozen or so QSPIRES MAIL files in WWW's rdr?

Received: by SLACVM (Mailier R2.08 R208004) id 2172;  
 Mon, 07 Feb 94 16:11:51 PST  
 Date: 7 Feb 1994, 16:11:51  
 From: WINTERS@SLACVM  
 To: BEBO  
 Subject: PROBTAK record 11398 updated.

Status has been changed from 1 to 6

Problem no : 11398  
 Prob Title : User Getting Partial SLAC Home Page  
 Entered by : WINTERS  
 Reported by : WINTERS <Joan M. Winters> (2530) []  
 Reported on : 02/03/94 At: 1450  
 Occurred on : 02/03/94 At: 1450  
 Problem Area: WW  
 Problem type: Support  
 Priority : 3 - Requires attention as time permits  
 Status : 6 - Bypassed, under investigation  
 Owner : BEBO

Description:

Bill Mayes (Mayes@unphys.phys.uh.edu), a Babar user, wrote:

I notice that to get the SLAC homepage, I have to  
 reload it many times. Otherwise it is cut off before the end of the page.  
 Also, I have only received the SLAC logo once a a complete item. Most  
 times I just get the top of the logo. I do not have hitisw trouble with any  
 other ww setup.

Within half an hour Mayes reported this problem to both Cottrell and me  
 that I know of.

When taken: (02/03/94 2304 WINTERS)

John, TonyJ, and I conferred on the phone about the problem. We think  
 probably another example of the problem of our SLACVM WW server  
 not correctly processing the HTTP 1.0 protocol. If so, Mayes's  
 problem is one that seems to have hit users of the Mac-for-Mosaic and  
 browsers, at least. I've written Mayes asking for more specifics  
 of his environment.

We think the problem is due to an interaction between the VM TCP/IP  
 and the HTTP daemon. If the HTTP daemon doesn't read all the  
 information sent from the WW browser to SLACVM's TCP/IP--and it may  
 be the client is using the HTTP 1.0 protocol (instead of the old  
 protocol)--and the HTTP daemon closes the socket while some  
 use data is still buffered in the TCP/IP DSM, then the TCP/IP DSM  
 is all its buffers for that process rather than sending off to the  
 browser the outgoing buffers it has. This is a documented  
 action/bug in VM TCP/IP.

It wouldn't be hard to modify the source to get our HTTPD  
 to read all the information sent by HTTP 1.0 clients. Right  
 now HTTPD only reads the first buffer, which is all HTTP 0.9 ever

(written up with John Halperin.)

Action taken: (02/04/94 2255 WINTERS)  
Bill Mayes wrote today that he's on a Macintosh Quadra 800, running  
NVA Mosaic, version 1.0.3 (27 Jan 1994). He also said that he's  
discovered he has this problem with other SIAC pages, too. He has not  
yet entered the problem elsewhere. I pointed him at the "Mosaic for  
Mac's HTTP 0.9 option. I asked him to let me know whether or not  
it bypasses his problem.

Action taken: (02/07/94 1603 WINTERS)  
Mail from Mayes today confirms that selecting the HTTP 0.9 option  
bypasses his page truncation problem.

ed: from SIACVM.BITNET by SIACVM.SIAC.STANFORD.EDU (Mailier R2.08 R208004)  
BSMTP Id 2795; Mon, 07 Feb 94 17:12:34 PST  
Mon, 7 Feb 1994 17:02:00 -0800  
"Joan M. Winters" <WINTERS@SIACVM.BITNET>  
"SIAC WorldWeb (WWW) Discussion List" <WWW-L@SIACVM.BITNET>  
Iterative Improvements  
Bebo White <BEO@SIACVM.BITNET>

- In light of Harvey's, also Pat's, feedback I've moved the Palo  
and Exploratorium links off the SIAC Home Page onto a new Local Area  
ces page.

planned to do this when I got a few more links. In light  
comments, I did it now. (I, too, especially didn't like having  
Resources before Other Experiments; but other feedback plus  
indicate that as the lists get longer, alphabetization is  
increasingly important as an aid to finding things.)

I think the local info is useful to SIAC workers, especially those new  
area, so I'm putting locres.html in the "Introductory" pages  
ry.

link to the page is near the end under Other Useful Info, Other  
Sources.

so the new SIAC logo (thanks to Terry Anderson, with an assist  
at). And the updated links to AIP FYI and Physics News Update.  
as to a pointer from John.) Plus the previous updates...  
Hastily, Joan

ed: by SLACVM (Mailier R2.08 R208004) Id 4168;

Wed, 09 Feb 1994 20:25 -0800 (PST)

<WINTERS@SLACVM>

: CC of Mail to Glanzman

Meant to carbon you on this but hit PFS too fast. Joan

[illegible]

Wed, 09 Feb 1994 19:58 -0800 (PST)

<WINTERS@SLACVM>

cc: Your WWW Mail of 3 Feb 1994 to Addis

On 3 Feb 1994 you said to Louise, who forwarded the onto MWM committee members:

MMM committee members:

we would find immediately useful

we would find immediately useful.

The ability to query an arbitrary oracle database

red your request onto Frank Rothacker and John Halperin  
om I've been talking about the need to get at Oracle from WMW.  
stand George Crane (who already got your note from Louise) is  
g about this, too.

It, you said:

we seem to be running version 1.1, while the SIAC documentation, incidentally, on such matters thinks it is version 1.2.

on such matters thinks it is version 1.2.

My place I know that our Mosaic version is documented in HTML, which does say 1.2. Is this your source?

HTWL, which does say 1.2. Is this your source?

to was written up by Bebo. Bebo, what's the true scoop?

Script You said:

(The minor exception concerned problems with PostScript non-portability)

problem with PostScript non-portability? Does this have to do with lack of support in `WWW/ghostscript` for Level 2 PostScript?

hear people are finding the web useful!

Joan

ed: by SLACVM (Mailier R2.08 R208004) Id 9898;

Tue, 08 Feb 94 12:21:38 PST

Tue, 08 Feb 1994 12:12 -0800 (PST)

"Joan M. Winters"

EBBO@SLACVM

cc: WW Updates?

I'm updating the WHATSNEW HTML page re page changes, which brought  
ed the WWSTAT page on system changes.

you made any major changes to SLAC's WW/Gopher system since your  
er 6 update? E.g., Mosaic for Mac? Is the production UNIX  
still CERN 2.12?

you update this info, would you also add the platforms for which  
with the vendor) supports the software? By the way, what do  
link of adding the support level for each of these software  
as, as defined in /usr/local/doc/policies/support-levels?

to have missed your Mosaic for Mac talk. I'd invited some ESnet  
to lunch since they were driving right past on their way to the  
meeting. How did the session go? Was there a lot of interest?  
Joan

ed: by SIACVM (Mailer R2.08 R208004) id 5835;

Wed, 09 Feb 94 09:02:54 PST

Wed, 09 Feb 1994 08:56 -0800 (PST)

"Frank Rothacker"

EB0@SIACVM

: Access control

ed the problem two years ago.

urns a path-id to the server. The server has to do a NETSTAT SOCKET  
id to find out who issued query.

easy when the client and server are both on VM. Otherwise you  
secure process running at the client IP address  
in supply the userid.

ed: by SLACVM (Mailier R2.08 R208004) Id 4446;  
Wed, 09 Feb 94 21:28:26 PST  
Wed, 09 Feb 1994 21:13 -0800 (PST)

<JXH@SLACVM>

"John Halperin"  
WINTERS@SLACVM, ADDIS@SLACVM  
DRAGON@SLACVM, BEBO@SLACVM  
t: Re: WWW/Oracle

passed on to me Tom's note about needing a WWW/Oracle  
ace. I know very little about Oracle, but just wanted to  
at that the same client/server wrapper approach we're using  
ix Qspires might work for sending individual queries from the  
WW server to an Oracle server, assuming only that Oracle  
es a dumb line-mode (ie, non-fullscreen, non-GUI) query  
ism. If Oracle provides a Perl or C language Application  
ace that would be even cleaner.

case, I would suggest that we strongly resist any temptation  
an Oracle interface to the VM WW server, even if it would  
y to do so with Oracle's Rexx interface (only available on  
far as I know).

's Mosaic version question: Bebo recently installed version  
with forms support) on at least AIX and Sun platforms.

Received: from UCSFVM.UCSF.EDU by SLACVM.SLAC.STANFORD.EDU (Mailier R2.08  
R208004) with BSMTP id 7076; Tue, 15 Feb 94 15:38:40 PST  
Received: from UCSFVM by UCSFVM.UCSF.EDU (Mailier R2.08 PTF008) with BSMTP id  
4863; Tue, 15 Feb 94 15:39:03 PST  
Received: from library.ucsf.edu by UCSFVM.UCSF.EDU (IBM VM SMTP V2R1) with TCP;  
Tue, 15 Feb 94 15:39:03 PST  
Received: from ora.com (ruby.ora.com) by library.ucsf.edu with SMTP id AA01845  
(5.67a8/IDA-1.5 for <sig-web@library.ucsf.edu>); Tue, 15 Feb 1994 15:07:02  
-0800

Received: from rock.west.ora.com by ruby (8.6.5/  
Received: by rock.west.ora.com (5.65c/Spike-2.1)  
id AA07248; Tue, 15 Feb 1994 15:09:15 -0800  
Message-Id: <199402152309.AA07248@rock.west.ora.com>  
The sig-web@library.ucsf.edu  
Subject: Volunteer meeting next week  
Date: Tue, 15 Feb 94 15:09:14 -0800  
From: "Christopher J. McKae" <mckae@ora.com>

Hi all,  
I'm busy pulling cable today and doing other stuff. We're not quite  
ready to meet yet, I think my message of last week just was too short  
notice. Let's plan on meeting Thursday evening of next week, exact time  
and location to be announced no later than this Thursday of this week.  
I'll be creating a mailing list for the volunteers. If you've  
already sent me a message or given me your card then you're already on  
the list. Otherwise, send a request to mckae@ora.com.  
Thanks for your commitment! See you soon.

Christopher McKae  
President, SIGWEB  
Project Manager  
Orally & Associates, Publishers  
mckae@ora.com

415/242-9623  
510/540-6036

Sent by SLACVM (Mailer R2.08 R208004) id 8975;

Mon, 14 Feb 94 20:39:08 PST

Mon, 14 Feb 1994 20:14 -0800 (PST)

John Halperin"

<JXH@SLACVM>

CRANE@SLACVM,

DDIS@slacvm.slac.stanford.edu,

lancey@slacvm.slac.stanford.edu

ebo@slacvm.slac.stanford.edu, tonyj@slacvm.slac.stanford.edu,

rane@slacvm.slac.stanford.edu, winters@slacvm.slac.stanford.edu,

alic@slacvm.slac.stanford.edu, pkreitz@slacvm.slac.stanford.edu,

cottrell@slacvm.slac.stanford.edu

Re: problem with WWW access to SLACVM

me and I restarted WWW and found it to still be unresponsive. The disk was 99% full, primarily because of the logging file. When I deleted the logging file and re-started the server everything was fine again. It should be easy enough to fill the disk up and reproduce the problem again if you like.

I know that (1) WWW's A disk has been full most of the times I checked for the last few weeks, yet the server still runs (at the time, at least); and (2) the console files from this time show that the server was busy handling requests after you restarted it, though apparently it was so busy with others' requests that it appeared unresponsive to you. I've appended a few lines of WWW's console file, starting from a few minutes before the first time you restarted it. The faint of heart might hit the discard key at about this point :-).

I'm like the server really was hung the first time, after the crash at 13:01:27, probably because the results could not be sent to the browser (maybe due to network problems, a bug in the browser, or the user's just stopping the browser; cf Probtrak for the general pathology.) There doesn't seem to have been anything wrong with the server the second time you restarted it other than just busy-ness.

Well, I'll bet you a soft drink of your choice that filling the disk will not reproduce the problem :-).

John

```

-----
Waiting for connection or message. (Mask=18 hex, max=6 hex).
waiting on socket 4
Reading socket 4 from host 128.200.29.6
Feb 14 12:59:26 1994 128.200.29.6 GET /FIND/spires?find+au+subbaraman
1994 12:59:26
128.200.29.6 FIND AU SUBBARAMAN Subfile= SPIRES
EXEC FSEARCH 128.200.29.6 find au subbaraman ( spires' returned 237
  
```

: Socket 4 disconnected by peer  
: Waiting for connection or message. (Mask=8 hex, max=6 hex).  
: New incoming connection:  
: Accepted new socket 4  
: Waiting for connection or message. (Mask=18 hex, max=6 hex).  
e waiting on socket 4  
: Reading socket 4 from host 134.93.178.136  
on Feb 14 12:59:43 1994 134.93.178.136 GET /FIND/spires?a+semenoff,g  
1994 12:59:43  
H 134.93.178.136 A SEMENOFF,G Subfile= SPIRES  
17  
d `EXEC FSEARCH 134.93.178.136 a semenoff,g ( spires' returned 5 lines.  
: Socket 4 disconnected by peer  
: Waiting for connection or message. (Mask=8 hex, max=6 hex).  
: New incoming connection:  
: Accepted new socket 4  
: Waiting for connection or message. (Mask=18 hex, max=6 hex).  
e waiting on socket 4  
: Reading socket 4 from host 134.79.16.22  
on Feb 14 12:59:47 1994 134.79.16.22 GET /FIND/abstracts  
8 FGET received: 134.79.16.22 ABSTRACTS (  
d `EXEC FGET 134.79.16.22 abstracts' returned 28 lines.  
: Socket 4 disconnected by peer  
: Waiting for connection or message. (Mask=8 hex, max=6 hex).  
: New incoming connection:  
: Accepted new socket 4  
: Waiting for connection or message. (Mask=18 hex, max=6 hex).  
e waiting on socket 4  
: Reading socket 4 from host 134.93.178.136  
on Feb 14 12:59:48 1994 134.93.178.136 GET /FIND/spires?a+semenoff,g  
1994 12:59:48  
I 134.93.178.136 A SEMENOFF,G Subfile= SPIRES  
10  
d `EXEC FSEARCH 134.93.178.136 a semenoff,g ( spires' returned 696  
: Socket 4 disconnected by peer  
: Waiting for connection or message. (Mask=8 hex, max=6 hex).  
: New incoming connection:  
: Accepted new socket 4  
: Waiting for connection or message. (Mask=18 hex, max=6 hex).  
e waiting on socket 4  
: Reading socket 4 from host 134.79.16.22  
on Feb 14 13:00:49 1994 134.79.16.22 GET /FIND/abstracts?fin+bull+9402020  
1994 13:00:49  
134.79.16.22 FIN BULL 9402020 Subfile= ABSTRACTS  
3  
d `EXEC FSEARCH 134.79.16.22 fin bull 9402020 ( abstracts' returned 21  
: Socket 4 disconnected by peer  
: Waiting for connection or message. (Mask=8 hex, max=6 hex).  
: New incoming connection:  
: Accepted new socket 4  
: Waiting for connection or message. (Mask=18 hex, max=6 hex).  
e waiting on socket 4  
: Reading socket 4 from host 128.135.102.20  
on Feb 14 13:00:53 1994 128.135.102.20 GET /FIND/spires  
1 FGET received: 128.135.102.20 SPIRES (

and `EXEC FGET 128.135.102.20 spires' returned 36 lines.  
 : Socket 4 disconnected by peer  
 : Waiting for connection or message. (Mask=8 hex, max=6 hex).  
 : New incoming connection:  
 : Accepted new socket 4  
 : Waiting for connection or message. (Mask=18 hex, max=6 hex).  
 : e waiting on socket 4  
 : Reading socket 4 from host 18.77.0.108  
 on Feb 14 13:00:54 1994 18.77.0.108 GET /FIND/SLAC.html HTTP/1.0  
 t: text/plain  
 t: application/x-html  
 t: application/html  
 t: text/x-html  
 t: text/html

t: application/x-hdf  
 t: application/x-netcdf  
 t: application/hdf  
 t: application/netcdf  
 4 FGET received: 18.77.0.108 SLAC.HTML (  
 and `EXEC FGET 18.77.0.108 SLAC.html' returned 269 lines.  
 : Socket 4 disconnected by peer  
 : Waiting for connection or message. (Mask=8 hex, max=6 hex).  
 : New incoming connection:  
 : Accepted new socket 4  
 : Waiting for connection or message. (Mask=18 hex, max=6 hex).  
 : e waiting on socket 4  
 : Reading socket 4 from host 128.84.219.115  
 on Feb 14 13:00:56 1994 128.84.219.115 GET  
 spires?find+author+bloom%2Ck HTTP/1.0

t: text/plain  
 t: application/x-html  
 t: application/html  
 t: text/x-html  
 t: text/html  
 : audio/\*  
 t: image/xwd  
 t: image/x-xwd  
 t: image/x-xwindowdump

1994 13:00:56

128.84.219.115 FIND AUTHOR BLOOM,K Subfile= SPIRES

Command `EXEC FSEARCH 128.84.219.115 find author bloom%2Ck ( spires'  
 11 lines.

: Socket 4 disconnected by peer  
 : Waiting for connection or message. (Mask=8 hex, max=6 hex).  
 : New incoming connection:  
 : Accepted new socket 4  
 : Waiting for connection or message. (Mask=18 hex, max=6 hex).  
 : e waiting on socket 4  
 : Reading socket 4 from host 18.77.0.108  
 on Feb 14 13:01:14 1994 18.77.0.108 GET /FIND/SLAC.html

4 FGET received: 18.77.0.108 SLAC.HTML (  
 d `EXEC FGET 18.77.0.108 SLAC.html' returned 269 lines.  
 : Socket 4 disconnected by peer  
 : Waiting for connection or message. (Mask=8 hex, max=6 hex).  
 : New incoming connection:

```

: Accepted new socket 4
: Waiting for connection or message. (Mask=18 hex, max=6 hex).
e waiting on socket 4
: Reading socket 4 from host 134.93.178.136
on Feb 14 13:01:16 1994 134.93.178.136 GET /FIND/spires?a+semenoff,g
1994 13:01:16
H 134.93.178.136 A SEMENOFF,G Subfile= SPIRES
27
nd `EXEC FSEARCH 134.93.178.136 a semenoff,g ( spires' returned 19 lines.
: Socket 4 disconnected by peer
: Waiting for connection or message. (Mask=8 hex, max=6 hex).
: New incoming connection:
: Accepted new socket 4
: Waiting for connection or message. (Mask=18 hex, max=6 hex).
e waiting on socket 4
: Reading socket 4 from host 128.84.219.115
on Feb 14 13:01:27 1994 128.84.219.115 GET
pires?find+author+bloom%2Ck
1994 13:01:27
H 128.84.219.115 FIND AUTHOR BLOOM,K Subfile= SPIRES
58
nd `EXEC FSEARCH 128.84.219.115 find author bloom%2Ck ( spires' returned
es.
E= 73:14:26 VIRTCPU= 013:17.07 TOTCPU= 020:03.20
AT 13:04:31 PST MONDAY 02/14/94 BY CRANE

```

```

ing WWW..
723I B (192) R/O
725I 192 also = D disk
723I E (193) R/O

```

Production system set

```

*****
*
* This disk contains C/370 Version 2 Release 1.0 *
*
* When using the software on this disk, be sure *
* that all other C/370 disks have been dropped. *
*
* Please direct questions or report problems to *
* Bebo White (BEBO@SLACVM) or phone to x2907. *
*
*****

```

( 197 G ) RO

```

*****
TCPAPPL DISK
/ IP for VM version V2R2 (5735-FAL 2.2)
routines, macros, and compilation execs for developing
TCP/IP applications on VM
programming interfaces are described in the IBM manual:
"TCP/IP for VM: Programmers Reference", SC31-6084-1.

```

\*\*\*\*\*

```
2 ( 198 H ) RO
S 13:05:07 PST MONDAY 02/14/94
T= 00:00:33 VIRTCPU= 000:00.33 TOTCPU= 000:00.48
: Parsed address as port 80, inet 0.0.0.0
: Opened socket number 3
: Master socket(), bind() and listen() all OK
: Waiting for connection or message. (Mask=8 hex, max=4 hex).
: New incoming connection:
: Accepted new socket 4
: Waiting for connection or message. (Mask=18 hex, max=5 hex).
: e waiting on socket 4
: Reading socket 4 from host 134.79.16.22
on Feb 14 13:05:11 1994 134.79.16.22 GET /FIND/slac.html
all FGET received: 134.79.16.22 SLAC.HTML (
and `EXEC FGET 134.79.16.22 slac.html' returned 269 lines.
: Socket 4 disconnected by peer
: Waiting for connection or message. (Mask=8 hex, max=5 hex).
: New incoming connection:
: Accepted new socket 4
: Waiting for connection or message. (Mask=18 hex, max=5 hex).
: e waiting on socket 4
: Reading socket 4 from host 134.93.178.136
107S Disk A(191) is full
on Feb 14 13:05:13 1994 134.93.178.136 GET /FIND/spires?a+semenoff,g
1994 13:05:13
H 134.93.178.136 A SEMENOFF,G Subfile= SPIRES
29
and `EXEC FSEARCH 134.93.178.136 a semenoff,g ( spires' returned 696
: Socket 4 disconnected by peer
: Waiting for connection or message. (Mask=8 hex, max=5 hex).
: New incoming connection:
: Accepted new socket 4
: Waiting for connection or message. (Mask=18 hex, max=5 hex).
: e waiting on socket 4
: Reading socket 4 from host 128.141.201.214
on Feb 14 13:05:40 1994 128.141.201.214 GET /FIND/SEMINARS/THISWEEK
0
t: www/source q=0.300
t: www/unknown q=0.300
t: application/octet-stream q=0.100
t: text/plain
t: text/html
t: www/source q=0.300
t: www/unknown q=0.300
41 SMGET received: 128.141.201.214 SEMINARS/THISWEEK (
t: application/ocCommand `EXEC FGET 128.141.201.214 SEMINARS/THISWEEK'
d 28 lines.
: Socket 4 disconnected by peer
: Waiting for connection or message. (Mask=8 hex, max=5 hex).
: New incoming connection:
: Accepted new socket 4
: Waiting for connection or message. (Mask=18 hex, max=5 hex).
: e waiting on socket 4
: Reading socket 4 from host 134.79.16.22
```

Mon Feb 14 13:05:43 1994 134.79.16.22 GET /FIND/abstracts  
43 FGET received: 134.79.16.22 ABSTRACTS (  
and `EXEC FGET 134.79.16.22 abstracts' returned 28 lines.  
n: Socket 4 disconnected by peer  
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).  
n: New incoming connection:  
n: Accepted new socket 4  
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).  
ge waiting on socket 4  
n: Reading socket 4 from host 134.93.178.136  
Mon Feb 14 13:05:43 1994 134.93.178.136 GET /FIND/spires?a+semenoff,g  
b 1994 13:05:43  
CH 134.93.178.136 A SEMENOFF,G Subfile= SPIRES  
50  
and `EXEC FSEARCH 134.93.178.136 a semenoff,g ( spires' returned 696  
n: Socket 4 disconnected by peer  
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).  
n: New incoming connection:  
n: Accepted new socket 4  
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).  
ge waiting on socket 4  
n: Reading socket 4 from host 128.138.242.17  
Mon Feb 14 13:07:37 1994 128.138.242.17 GET /FIND/SLAC.html HTTP/1.0  
ot: text/plain  
ot: application/x-html  
ot: application/html  
ot: text/x-html  
ot: text/html  
ot: audio/basic  
ot: audio/x-aiff  
ot: image/gif  
ot: image/jpeg  
ot: image/tiff  
37 FGET received: 128.138.242.17 SLAC.HTML (  
ot: Command `EXEC FGET 128.138.242.17 SLAC.html' returned 269 lines.  
n: Socket 4 disconnected by peer  
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).  
n: New incoming connection:  
n: Accepted new socket 4  
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).  
ge waiting on socket 4  
n: Reading socket 4 from host 128.141.201.214  
Mon Feb 14 13:07:39 1994 128.141.201.214 GET /FIND/SEMINARS/TOMORROW  
0  
t: www/source q=0.300  
t: www/unknown q=0.300  
t: application/octet-stream q=0.100  
t: text/plain  
t: text/html  
t: www/source q=0.300  
t: www/unknown q=0.300  
40 SMGET received: 128.141.201.214 SEMINARS/TOMORROW (  
t: application/ocCommand `EXEC FGET 128.141.201.214 SEMINARS/TOMORROW'  
d 46 lines.  
n: Socket 4 disconnected by peer  
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).

n: New incoming connection:  
n: Accepted new socket 4  
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).  
ge waiting on socket 4  
n: Reading socket 4 from host 130.199.80.4  
Mon Feb 14 13:07:47 1994 130.199.80.4 GET /FIND/apsnews.html HTTP/1.0  
pt: text/plain  
pt: application/x-html  
pt: application/html  
pt: text/x-html  
pt: text/html  
pt: application/x-hdf  
pt: application/x-netcdf  
pt: application/hdf  
:47 FGET received: 130.199.80.4 APSNEWS.HTML (  
pt: application/netcdfCommand `EXEC FGET 130.199.80.4 apsnews.html'  
ed 131 lines.  
n: Socket 4 disconnected by peer  
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).  
n: New incoming connection:  
n: Accepted new socket 4  
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).  
ge waiting on socket 4  
n: Reading socket 4 from host 128.138.242.17  
Mon Feb 14 13:07:48 1994 128.138.242.17 GET /FIND/SLAC.html  
:48 FGET received: 128.138.242.17 SLAC.HTML (  
and `EXEC FGET 128.138.242.17 SLAC.html' returned 269 lines.  
n: Socket 4 disconnected by peer  
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).  
n: New incoming connection:  
n: Accepted new socket 4  
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).  
ge waiting on socket 4  
n: Reading socket 4 from host 130.199.80.4  
Mon Feb 14 13:07:50 1994 130.199.80.4 GET /FIND/apsnews.html  
:50 FGET received: 130.199.80.4 APSNEWS.HTML (  
and `EXEC FGET 130.199.80.4 apsnews.html' returned 131 lines.  
n: Socket 4 disconnected by peer  
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).  
n: New incoming connection:  
n: Accepted new socket 4  
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).  
ge waiting on socket 4  
n: Reading socket 4 from host 128.138.242.17  
Mon Feb 14 13:07:51 1994 128.138.242.17 GET /FIND/slac.xbm HTTP/1.0  
pt: text/plain  
pt: application/x-html  
pt: application/html  
pt: text/x-html  
pt: text/html  
pt: audio/basic  
pt: audio/x-aiff  
pt: image/gif  
pt: image/jpeg  
pt: image/tiff

```

13:52 FGET received: 128.138.242.17 SLAC.XBM (
pt: imCommand `EXEC FGET 128.138.242.17 slac.xbm' returned 195 lines.
n: Socket 4 disconnected by peer
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).
n: New incoming connection:
n: Accepted new socket 4
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).
ge waiting on socket 4
n: Reading socket 4 from host 134.93.178.136
Mon Feb 14 13:07:53 1994 134.93.178.136 GET /FIND/spires?a+semenoff,g
b 1994 13:07:53
CH 134.93.178.136 A SEMENOFF,G Subfile= SPIRES
13:09
and `EXEC FSEARCH 134.93.178.136 a semenoff,g ( spires' returned 696
n: Socket 4 disconnected by peer
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).
n: New incoming connection:
n: Accepted new socket 4
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).
ge waiting on socket 4
n: Reading socket 4 from host 128.138.242.17
Mon Feb 14 13:08:27 1994 128.138.242.17 GET /FIND/slac.xbm

13:28 FGET received: 128.138.242.17 SLAC.XBM (
and `EXEC FGET 128.138.242.17 slac.xbm' returned 195 lines.
n: Socket 4 disconnected by peer
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).
n: New incoming connection:
n: Accepted new socket 4
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).
ge waiting on socket 4
n: Reading socket 4 from host 130.199.80.4
Mon Feb 14 13:08:29 1994 130.199.80.4 GET /FIND/an931210.html HTTP/1.0
pt: text/plain
pt: application/x-html
pt: application/html
pt: text/x-html
pt: text/html
pt: application/x-hdf
pt: application/x-netcdf
pt: application/hdf
13:29 FGET received: 130.199.80.4 AN931210.HTML (
pt: application/netcdCommand `EXEC FGET 130.199.80.4 an931210.html'
ed 71 lines.
n: Socket 4 disconnected by peer
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).
n: New incoming connection:
n: Accepted new socket 4
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).
ge waiting on socket 4
n: Reading socket 4 from host 128.165.23.9
Mon Feb 14 13:08:30 1994 128.165.23.9 GET
abstracts?fin+bull+hep-lat+9211061
b 1994 13:08:30
CH 128.165.23.9 FIN BULL HEP-LAT 9211061 Subfile= ABSTRACTS
13:39

```

Command `EXEC FSEARCH 128.165.23.9 fin bull hep-lat 9211061 ( abstracts'  
ed 13 lines.  
n: Socket 4 disconnected by peer  
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).  
n: New incoming connection:  
n: Accepted new socket 4  
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).  
ge waiting on socket 4  
n: Reading socket 4 from host 134.93.178.136  
Mon Feb 14 13:08:41 1994 134.93.178.136 GET /FIND/spires?a+semenoff,g  
Feb 1994 13:08:42  
RCH 134.93.178.136 A SEMENOFF,G Subfile= SPIRES  
3:55  
Command `EXEC FSEARCH 134.93.178.136 a semenoff,g ( spires' returned 696  
n: Socket 4 disconnected by peer  
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).  
n: New incoming connection:  
n: Accepted new socket 4  
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).  
ge waiting on socket 4  
n: Reading socket 4 from host 143.202.88.48  
Mon Feb 14 13:09:08 1994 143.202.88.48 GET /FIND/slac.xbm HTTP/1.0  
pt: text/plain  
pt: application/x-html  
pt: application/html  
pt: text/x-html  
pt: text/html  
pt: application/x-hdf  
pt: application/x-netcdf  
pt: application/hdf  
pt: application/netcdf  
:08 FGET received: 143.202.88.48 SLAC.XBM (  
Command `EXEC FGET 143.202.88.48 slac.xbm' returned 195 lines.  
n: Socket 4 disconnected by peer  
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).  
n: New incoming connection:  
n: Accepted new socket 4  
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).  
ge waiting on socket 4  
n: Reading socket 4 from host 128.141.201.214  
Mon Feb 14 13:09:09 1994 128.141.201.214 GET /FIND/SLAC.html HTTP/1.0  
pt: www/source q=0.300  
pt: www/unknown q=0.300  
pt: application/octet-stream q=0.100  
pt: text/plain  
pt: text/html  
pt: www/source q=0.300  
pt: www/unknown q=0.300  
:10 FGET received: 128.141.201.214 SLAC.HTML (  
pt: application/octet-streCommand `EXEC FGET 128.141.201.214 SLAC.html'  
ed 269 lines.  
n: Socket 4 disconnected by peer  
n: Waiting for connection or message. (Mask=8 hex, max=5 hex).  
n: New incoming connection:  
n: Accepted new socket 4  
n: Waiting for connection or message. (Mask=18 hex, max=5 hex).

ge waiting on socket 4

n: Reading socket 4 from host 143.202.88.48

Mon Feb 14 13:09:11 1994 143.202.88.48 GET /FIND/slac.xbm

9:12 FGET received: 143.202.88.48 SLAC.XBM (  
and `EXEC FGET 143.202.88.48 slac.xbm' returned 195 lines.

n: Socket 4 disconnected by peer

n: Waiting for connection or message. (Mask=8 hex, max=5 hex).

n: New incoming connection:

n: Accepted new socket 4

n: Waiting for connection or message. (Mask=18 hex, max=5 hex).

ge waiting on socket 4

n: Reading socket 4 from host 134.93.178.136

Mon Feb 14 13:09:16 1994 134.93.178.136 GET /FIND/spires?a+semenoff,g

b 1994 13:09:16

CH 134.93.178.136 A SEMENOFF,G Subfile= SPIRES

27

and `EXEC FSEARCH 134.93.178.136 a semenoff,g ( spires' returned 696

G - 07:30:01 PST MONDAY 02/14/94

### System Messages.

0027 RDR, 0001 PRT, NO PUN

NECTED AT 13:09:59 PST MONDAY 02/14/94

MS

A Rel. 1.1 10/15/93 10:12

; T=0.16/0.24 13:10:03

; T=0.01/0.01 13:10:10

ve store daemon log

(S) NOT FOUND.

(00028); T=0.10/0.12 13:10:16

N LOG A1

MON @MODULE A1

MON MODULE A1

NG GLOBALV A0

LE EXEC A2

EMO EXEC A1

SC EXEC A1

DIRECT A0

; T=0.01/0.01 13:10:18

ve store daemon log

FILE 2301 SENT TO ARCHIVES RDR AS 2808 RECS 059K CPY 001 B NOHOLD

; T=0.84/3.56 13:11:47

daemon log a

; T=0.01/0.02 13:11:52

CT= 00:07:20 VIRTCPU= 000:04.71 TOTCPU= 000:09.41

F AT 13:11:54 PST MONDAY 02/14/94