A Repository with 44 Years of Unix Evolution

github.com/dspinellis/unix-history-make

Diomidis Spinellis
Department of Management Science and Technology
Athens University of Economics and Business
www.spinellis.gr
@CoolSWEng
dds@aueb.gr

PRIMARY MATERIALS (8.4GB)

Index of /Archive/PDP-11/Dist

MOST MATERIAL CAME FROM:
THE UNIX HISTORY SOCIETY SITE (TUHS.ORG),
THE BERKELEY CSRG CD-ROMS,
THE OLD LINUX SITE (OLDLINUX.ORG), AND
THE FREEBSD ARCHIVES.

HOW YOU CAN HELP

- USE THE REPO FOR YOUR RESEARCH
- ADD CODE AUTHORSHIP INFO
- IMPORT 2.* BSD, NETBSD, OPENBSD
- LOBBY TO FREE SYSTEM V SOURCE CODE

A GROUP HAS OCR'D FIRST EDITION UNIX FROM PRINTOUTS

Here's why Digital's PDP-11/70 is the fastest selling supermini in the world.

PDP-7

BELL LABS

TELETYPE MODEL 33

VT-52

DENNIS RITCHIE & KEN THOMPSON ON A PDP-11

DEC VAX

VT-100

386 BSD REFERENCE CD-ROM — WILLIAM & LYNNE JOLITZ

COMMIT EXAMPLES

1970s: THE BELL LABS TEAM

```
COMMIT 94A21182365EBB258EEEE2AA41C5FBCB1F7FD566
\ AUTHOR: \ KEN \ THOMPSON AND \ DENNIS \ RITCHIE <RESEARCH!(KEN,DMR)>
\ DATE:   \ TUE \ JUN 20 04:00:00 1972 -0500
    \ RESEARCH \ V1 DEVELOPMENT
        \ WORK ON FILE U5.S
[...]

COMMIT B7B2640B9E27415D453A8FBE9754879C
\ AUTHOR: \ KEN \ THOMPSON <RESEARCH!KEN>
\ DATE:   \ TUE \ NOV 26 18:13:21 1974 -05
    \ RESEARCH \ V5 DEVELOPMENT
        \ WORK ON FILE USR/SYS/KEN/SLP.C
[...]

COMMIT 3D19667A65D35A411D911282ED8887E32
\ AUTHOR: \ DENNIS \ RITCHIE <RESEARCH!DMR>
\ DATE:   \ MON \ DEC 2 18:18:02 1974 -0500
    \ RESEARCH \ V5 DEVELOPMENT
        \ WORK ON FILE USR/SYS/DMR/KL.C

COMMIT 171931A3F6F28CE4D196C20FDEC6A4413A843B95
\ AUTHOR: \ BRIAN \ W. \ KERNIGHAN <RESEARCH!BWK>
\ DATE:   \ TUE \ MAY 13 19:43:47 1975 -0500
    \ RESEARCH \ V6 DEVELOPMENT
        \ WORK ON FILE RAT/R.G
[...]

COMMIT AC4B13BCA433A44A97689AF1024797011883469
\ AUTHOR: \S. \R. \BOURNE <RESEARCH!SRB>
\ DATE:   \ FRI \ JAN 12 02:17:45 1979 -05
    \ RESEARCH \ V7 DEVELOPMENT
        \ WORK ON FILE USR/SRC/CMD/SH/BLOK.C
```

NATIONAL MEDAL OF TECHNOLOGY AWARD

THE C PROGRAMMING LANGUAGE

STEPHEN R. BOURNE WROTE THE SHELL BEARING HIS NAME

1980s: UNIX GOES TO BERKELEY

Sun microsystems

```
commit a20c66bfa7223fbaf148f13b47356ad4326e9265
Author: Bill Joy <wnj@ucbvax.Berkeley.EDU>
Date:   Sun Oct 24 22:38:00 1982 -0800

    create netif directory

    SCCS-vsn: 4.5

[...]

commit e449c1d11a5de86f09ea0c5b80461d7d6cb01792
Author: Eric Schmidt <schmidt@ucbvax.Berkeley.EDU>
Date:   Wed May 9 18:47:25 1979 -0800

    BSD 2 development

    Work on file src/net/netcp.c
```

BILL JOY: CO-FOUNDER OF SUN MICROSYSTEMS

ERIC SCHMIDT, NOW EXECUTIVE CHAIRMAN OF GOOGLE

Google

ERIC ALLMAN COMMITS THE SENDMAIL DEBUG MODE THAT WILL BE USED BY THE FIRST INTERNET WORM

```
commit a7436288bb7f734dfe0b44d3c5d58f9d57fbb6c0
Author: Eric Allman <eric@ucbvax.Berkeley.EDU>
Date:   Sat Jun 26 02:38:29 1982 -0800

    add debugging information to server smtp code: _debug
    Debug, _verbose to set verbose, and _showq to show the queue

    SCCS-vsn: 3.18

index 1ec3c41..1eadb47 100644
--- a/usr/src/usr.sbin/sendmail/src/srvrsmtp.c
+++ b/usr/src/usr.sbin/sendmail/src/srvrsmtp.c
@@ -55,7 +57,9 @@ static struct cmd   CmdTab[] =
        "helo",     CMDHELO,
 # ifdef DEBUG
-       "showq",    CMDDBGSHOWQ,
+       "_showq",   CMDDBGSHOWQ,
+       "_debug",   CMDDBGDEBUG,
+       "_verbose", CMDDBGVERBOSE,
 # endif DEBUG
        NULL,       CMDERROR,
@@ -245,6 +252,16 @@ smtp()
        break;
+
+       case CMDDBGDEBUG:   /* set debug mode */
+           Debug = atoi(p);
+           message("200", "Debug = %d", Debug);
+           break;
# endif DEBUG
```

COMMUNICATIONS acm

REPOSITORY STRUCTURE AND CONTENTS

```
V1 → V3 → V4 → V5 → V6 → V7
```

IMPORTED UNIX SNAPSHOTS, REPOSITORIES, AND THEIR MERGERS

1972 First Edition:
4,768 lines
of PDP-11 code

Bell Laboratories

```
BSD 1
  ↓
BSD 2 → Unix 32/V
  ↓
BSD 3
  ↓
BSD SCCS
```

Creation Process

- Populate maps of files to authors
- Fill-in author details
- Import release snapshots
- Import CSRC SCCS repo
- Import FreeBSD git repos

```
BSD 4.0  BSD 4.1  BSD 4.2  BSD 4.3  BSD 4.3-Tahoe  BSD 4.3-Reno  BSD 4.4
                             BSD 4.3 Net/1  BSD 4.3 Net/2  BSD 4.4 Lite1
                             386BSD 0.0                    BSD 4.4 Lite2
                             386BSD 0.1
                             FreeBSD 1 CVS
                             FreeBSD Git
```

Commits per year

- 659 thousand commits
- 2306 merges
- 850 code authors
- 24 imported snapshots
- 40 tags
- 149 branches

Snapshot import
VCS import

Bell Labs   Berkeley   Open source (including Berkeley)

386 BSD

2015 FreeBSD 10: 26 million lines of code

FreeBSD

CODE PROVENANCE ACROSS SIGNIFICANT UNIX RELEASES

Code provenance
F9 — FreeBSD 9.0
F8 — FreeBSD 8.0
FS2 — FreeBSD 5.2
F4 — FreeBSD 4.0
F2 — FreeBSD 2.0
F1 — FreeBSD 1.0
3B1 — 386BSD 0.1
BN2 — BSD 4.3 Net/2
B43 — BSD 4.3
B42 — BSD 4.2
B41 — BSD 4.1
V7 — Research Version 7
V6 — Research Version 6

Lines of code (millions)

Unix release: V6 V7 B41 B42 B43 BN2 3B1 F1 F2 F4 FS2 F8 F9

```
$ git blame -C -C usr/sys/sys/pipe.c
3cc1108b usr/sys/sys/pipe.c  (Ken Thompson 1974-11-26 18:13:21 -0500 53)  rf->f_flag = FREAD|FPIPE;
3cc1108b usr/sys/sys/pipe.c  (Ken Thompson 1974-11-26 18:13:21 -0500 54)  rf->f_inode = ip;
3cc1108b usr/sys/sys/pipe.c  (Ken Thompson 1974-11-26 18:13:21 -0500 55)  ip->i_count = 2;
[...]
1f183be2 usr/sys/sys/pipe.c  (Ken Thompson 1979-01-10 15:19:35 -0500 122)
1f183be2 usr/sys/sys/pipe.c  (Ken Thompson 1979-01-10 15:19:35 -0500 123)  register struct inode *ip;
1f183be2 usr/sys/sys/pipe.c  (Ken Thompson 1979-01-10 15:19:35 -0500 124)
1f183be2 usr/sys/sys/pipe.c  (Ken Thompson 1979-01-10 15:19:35 -0500 125)  ip = fp->f_inode;
1f183be2 usr/sys/sys/pipe.c  (Ken Thompson 1979-01-10 15:19:35 -0500 126)  c = u.u_count;
1f183be2 usr/sys/sys/pipe.c  (Ken Thompson 1979-01-10 15:19:35 -0500 127) loop:
1f183be2 usr/sys/sys/pipe.c  (Ken Thompson 1979-01-10 15:19:35 -0500 128)  /*
1f183be2 usr/sys/sys/pipe.c  (Ken Thompson 1979-01-10 15:19:35 -0500 129)   * If error or all done, return.
9a9f6b22 usr/src/sys/sys/pipe.c  (Bill Joy 1980-01-05 05:51:18 -0800 130)   */
9a9f6b22 usr/src/sys/sys/pipe.c  (Bill Joy 1980-01-05 05:51:18 -0800 131)
9a9f6b22 usr/src/sys/sys/pipe.c  (Bill Joy 1980-01-05 05:51:18 -0800 132)  if (u.u_error)
9a9f6b22 usr/src/sys/sys/pipe.c  (Bill Joy 1980-01-05 05:51:18 -0800 133)      return;
9a9f6b22 usr/src/sys/sys/pipe.c  (Bill Joy 1980-01-05 05:51:18 -0800 134)  plock(ip);
6d632e85 usr/sys/sys/pipe.c  (Ken Thompson 1975-07-17 10:33:37 -0500 135)  if(c == 0) {
6d632e85 usr/sys/sys/pipe.c  (Ken Thompson 1975-07-17 10:33:37 -0500 136)      prele(ip);
6d632e85 usr/sys/sys/pipe.c  (Ken Thompson 1975-07-17 10:33:37 -0500 137)      u.u_count = 0;
6d632e85 usr/sys/sys/pipe.c  (Ken Thompson 1975-07-17 10:33:37 -0500 138)      return;
6d632e85 usr/sys/sys/pipe.c  (Ken Thompson 1975-07-17 10:33:37 -0500 139)  }
6d632e85 usr/sys/sys/pipe.c  (Ken Thompson 1975-07-17 10:33:37 -0500 140) }

lib/libc/gen/timezone.c:975fa769 usr/src/lib/gen/timezone.c (<research!dmr> 1979-01-10 14:58:45 -0500 76) static struct zone {
lib/libc/gen/timezone.c:975fa769 usr/src/lib/gen/timezone.c (<research!dmr> 1979-01-10 14:58:45 -0500 77)   int   offset;
lib/libc/gen/timezone.c:975fa769 usr/src/lib/gen/timezone.c (<research!dmr> 1979-01-10 14:58:45 -0500 78)   char  *stdzone;
lib/libc/gen/timezone.c:975fa769 usr/src/lib/gen/timezone.c (<research!dmr> 1979-01-10 14:58:45 -0500 79)   char  *dlzone;
lib/libc/gen/timezone.c:975fa769 usr/src/lib/gen/timezone.c (<research!dmr> 1979-01-10 14:58:45 -0500 80) } zonetab[] = {
lib/libc/gen/timezone.c:975fa769 usr/src/lib/gen/timezone.c (<research!dmr> 1979-01-10 14:58:45 -0500 113)   if (dst && zp->dlzone)
lib/libc/gen/timezone.c:975fa769 usr/src/lib/gen/timezone.c (<research!dmr> 1979-01-10 14:58:45 -0500 114)       return(zp->dlzone);
lib/libc/gen/timezone.c:975fa769 usr/src/lib/gen/timezone.c (<research!dmr> 1979-01-10 14:58:45 -0500 115)   if (!dst && zp->stdzone)
lib/libc/gen/timezone.c:975fa769 usr/src/lib/gen/timezone.c (<research!dmr> 1979-01-10 14:58:45 -0500 116)       return(zp->stdzone);
```

GIT BLAME CORRECTLY ATTRIBUTES AUTHORS AND PROVIDES COMMIT DATES

THE OLDEST SURVIVING CODE IN MODERN FREEBSD UNIX

CODE STYLE EVOLUTION ALONG UNIX RELEASES

- Mean filename length
- Mean identifier length
- Comments per 100 lines
- % of goto statements
- "register" keywords per 100 lines

Value (see legend)

Mean date of release's files

LIVE FREE OR DIE

UNIX