

---

# The Unix Toolchest: Power Tools for Power Programmers

---

**Diomidis Spinellis**

[www.spinellis.gr](http://www.spinellis.gr)

CBS



I run  
Microsoft  
Office



So do I



See the latest episodes of PC and Mac  
[apple.com/uk](http://apple.com/uk)



25

Effective SOFTWARE DEVELOPMENT SERIES   
Scott Meyers, Consulting Editor

# CODE *Reading*

*The Open Source Perspective*



Diomidis Spinellis

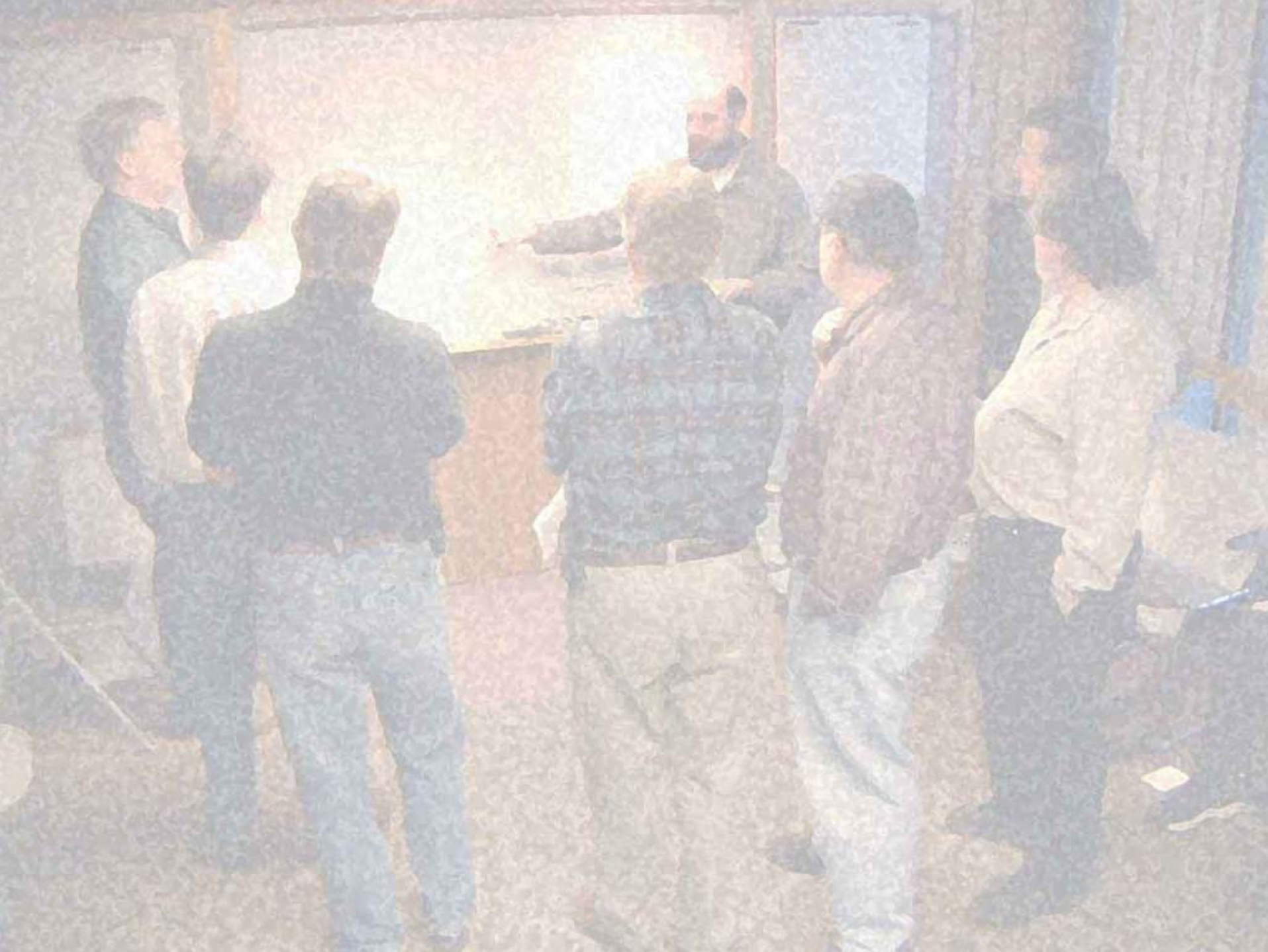
Effective SOFTWARE DEVELOPMENT SERIES   
Scott Meyers, Consulting Editor

# CODE *Quality*



Diomidis Spinellis

Advantages







Efficient



digital pdp11

digital pdp11

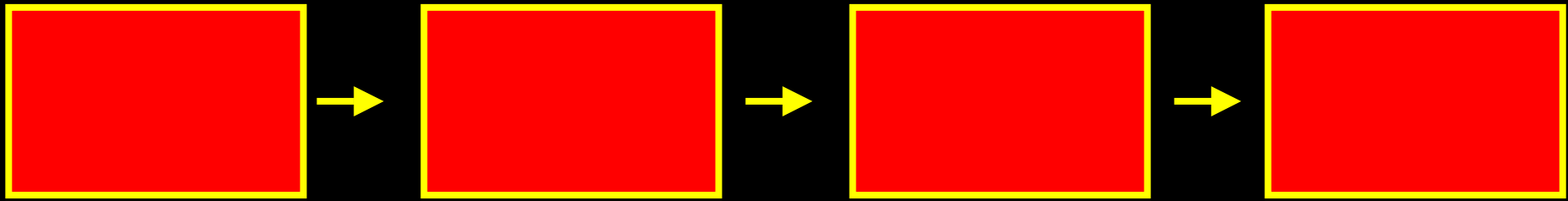
pdp11  
digital equipment corporation

000000

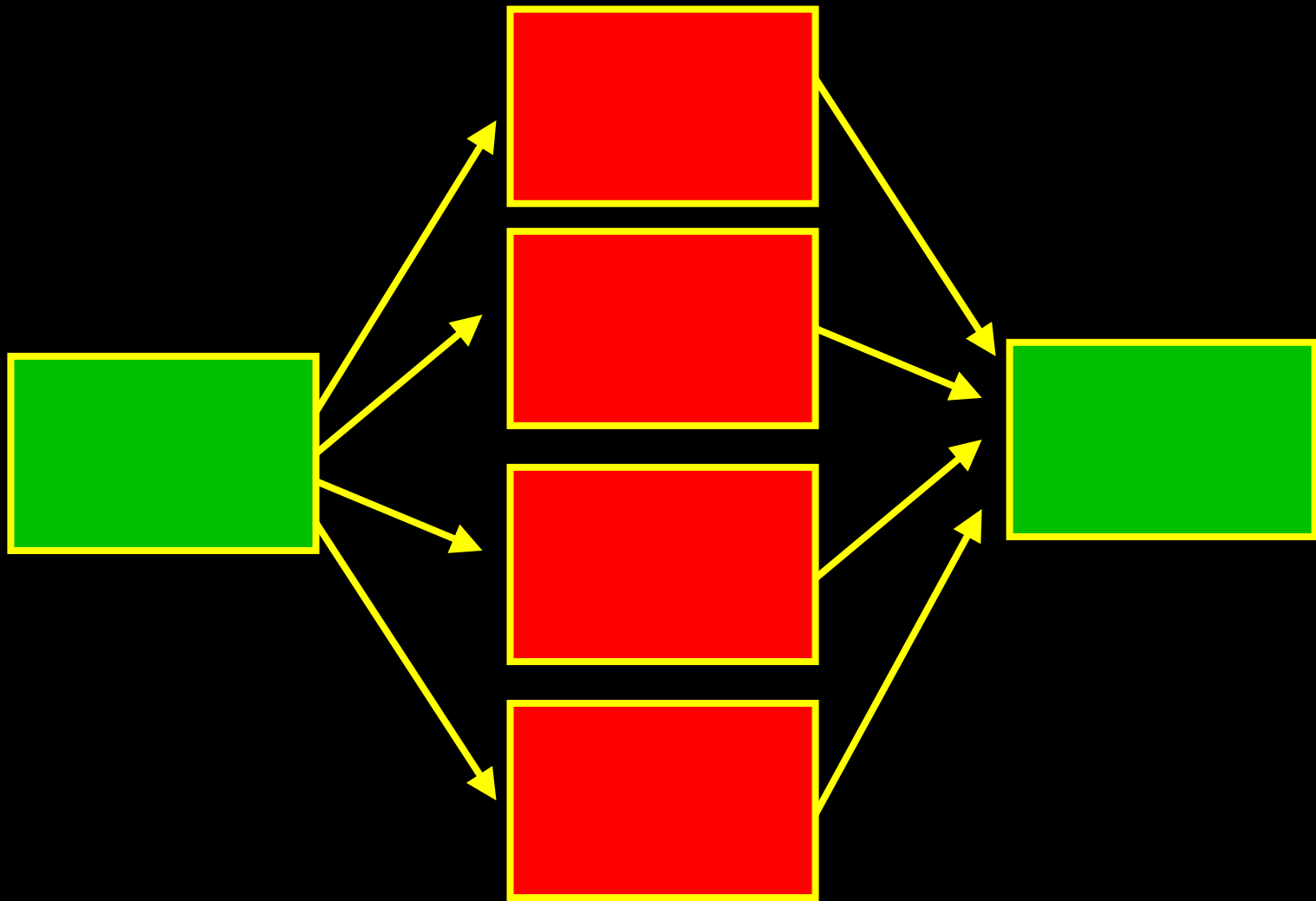
pdp11







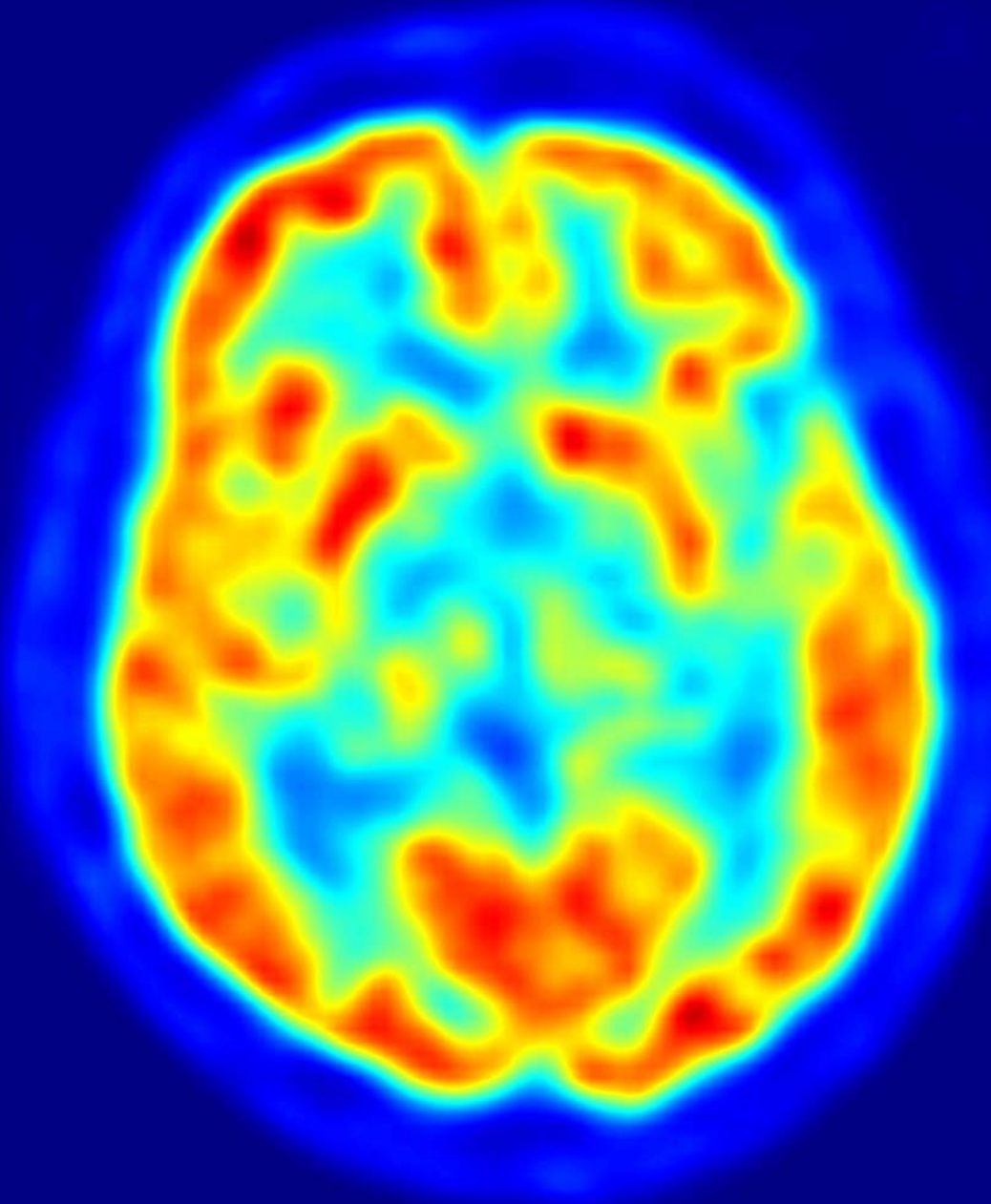
interleaving













**open source**



small is beautiful

```
main(int argc, char *argv[])
{
    while (*argv) {
        (void)printf("%s", *argv);
        if (*++argv)
            putchar(' ');
    }
    putchar('\n');
    exit(0);
}
```

```
# $FreeBSD: src/etc/master.passwd,v 1.40 2005/06/06 20:19:56 brooks Exp $
#
root:*:0:0:Charlie &:/root:/bin/csh
toor:*:0:0:Bourne-again Superuser:/root:
daemon*:1:1:Owner of many system processes:/root:/usr/sbin/nologin
operator*:2:5:System &:/usr/sbin/nologin
bin*:3:7:Binaries Commands and Source:/usr/sbin/nologin
tty*:4:65533:Tty Sandbox:/usr/sbin/nologin
mem*:5:65533:KMem Sandbox:/usr/sbin/nologin
games*:7:13:Games pseudo-user:/usr/games:/usr/sbin/nologin
news*:8:8:News Subsystem:/usr/sbin/nologin
man*:9:9:Mister Man Pages:/usr/share/man:/usr/sbin/nologin
sshd*:22:22:Secure Shell Daemon:/var/empty:/usr/sbin/nologin
smmisp*:25:25:Sendmail Submission
User:/var/spool/clientmqueue:/usr/sbin/nologin
mailnull*:26:26:Sendmail Default
User:/var/spool/mqueue:/usr/sbin/nologin
bind*:53:53:Bind Sandbox:/usr/sbin/nologin
proxy*:62:62:Packet Filter pseudo-user:/nonexistent:/usr/sbin/nologin
_pflodg*:64:64:pflodg privsep user:/var/empty:/usr/sbin/nologin
_dhcp*:65:65:dhcp programs:/var/empty:/usr/sbin/nologin
uucp*:66:66:UUCP pseudo-
user:/var/spool/uucppublic:/usr/local/libexec/uucp/uucico
pop*:68:6:Post Office Owner:/nonexistent:/usr/sbin/nologin
www*:80:80:World Wide Web Owner:/nonexistent:/usr/sbin/nologin
```





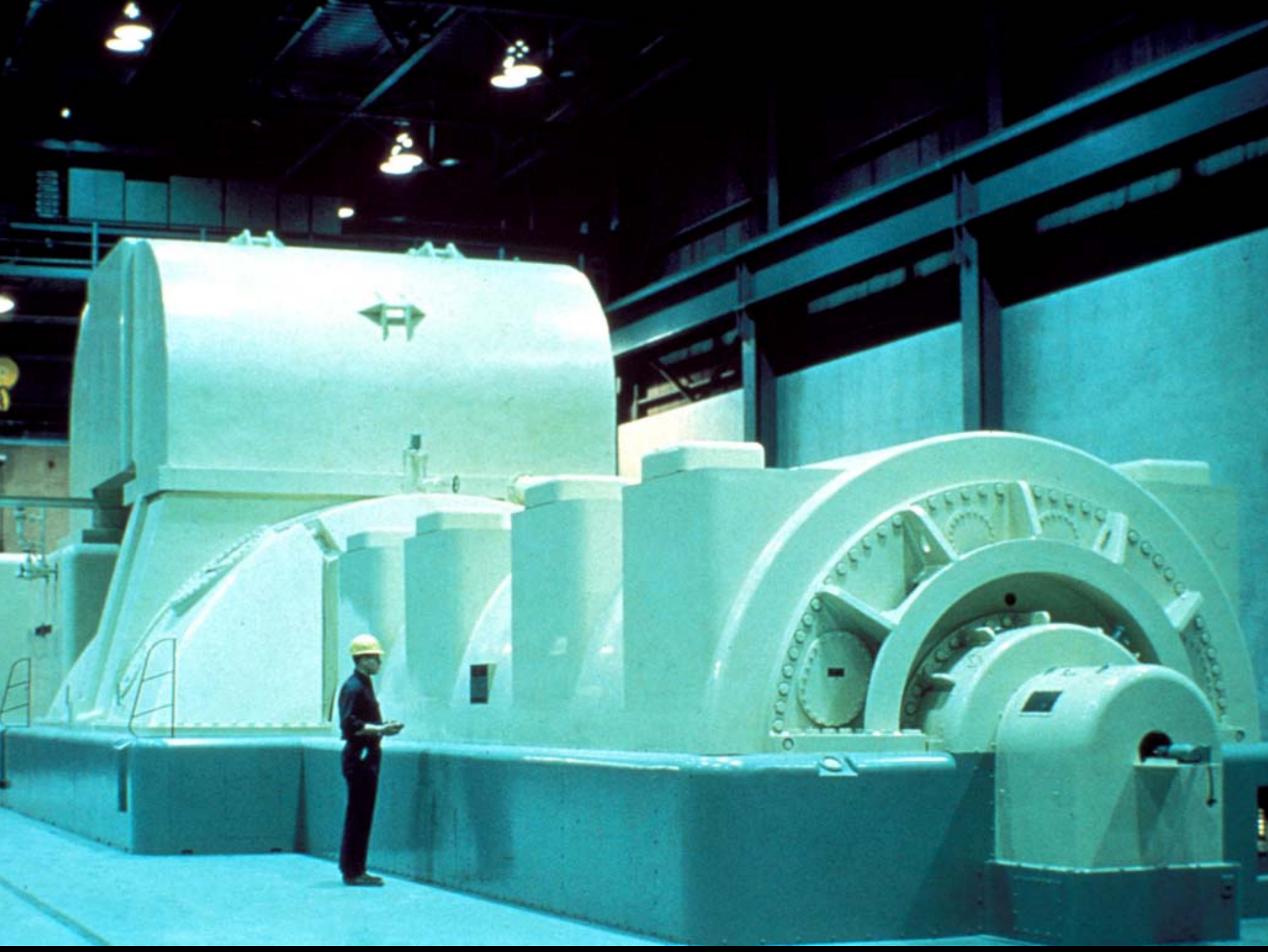




Get



Generate



Select





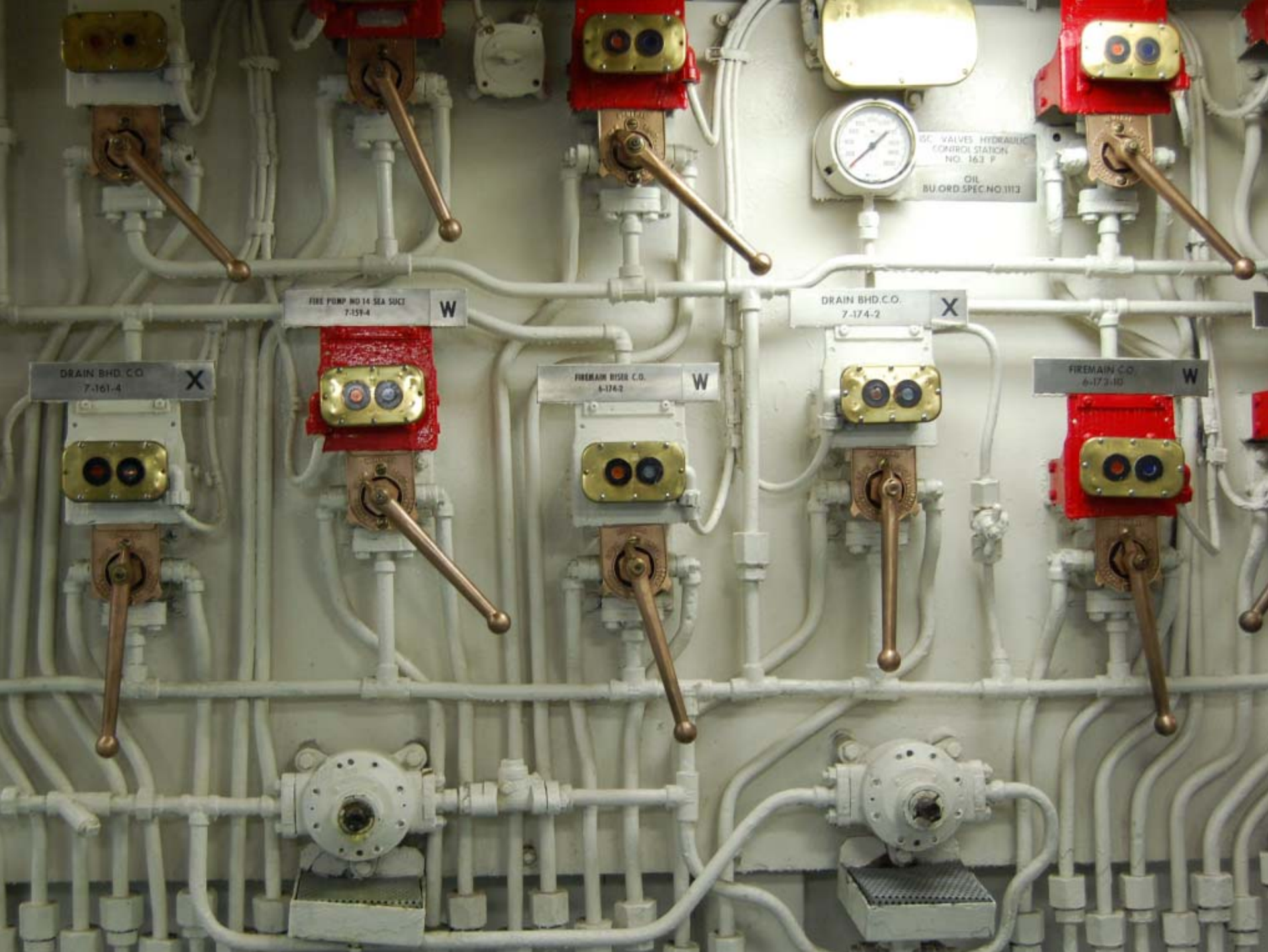
Process



Summarize



Plumb



G.C. VALVES HYDRAULIC  
CONTROL STATION  
NO. 163 P  
OIL  
BU ORD SPEC NO 1113

FIRE PUMP NO 14 SEA SUCT  
7-159-4 W

DRAIN BHD. CO.  
7-174-2 X

DRAIN BHD. CO.  
7-161-4 X

FIREMAN RISER C.O.  
6-174-2 W

FIREMAN C.O.  
6-173-10 W

Program







curl

wget

```
$ while read ticker
do
    curl -o $ticker.html
"http://www.reuters.com
/finance/stocks/ratios?
symbol=$ticker&rpc=66#m
anagement"
done <ticker_symbols
```

mysql

sqlplus

osql

sqlite3

psql

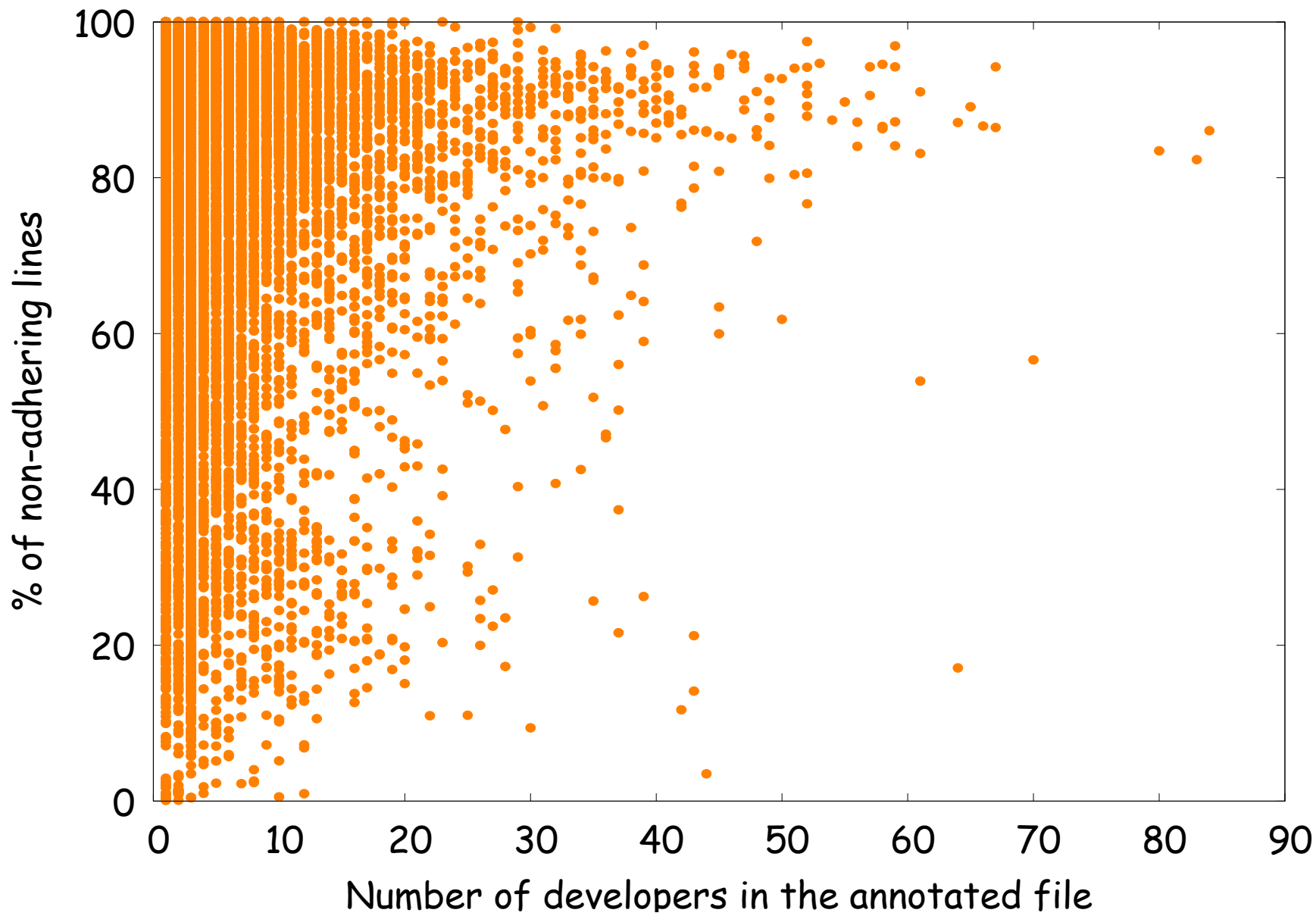
odbc

```
$ mysql -s db -e "select
distinct substr(name, 1,
length(name) -
instr(reverse(name), '/'))
from FILES" |
xargs ls -di |
awk '{print $1}' |
sort -u |
wc -l
```

**svn / cvs log**

**svn blame**

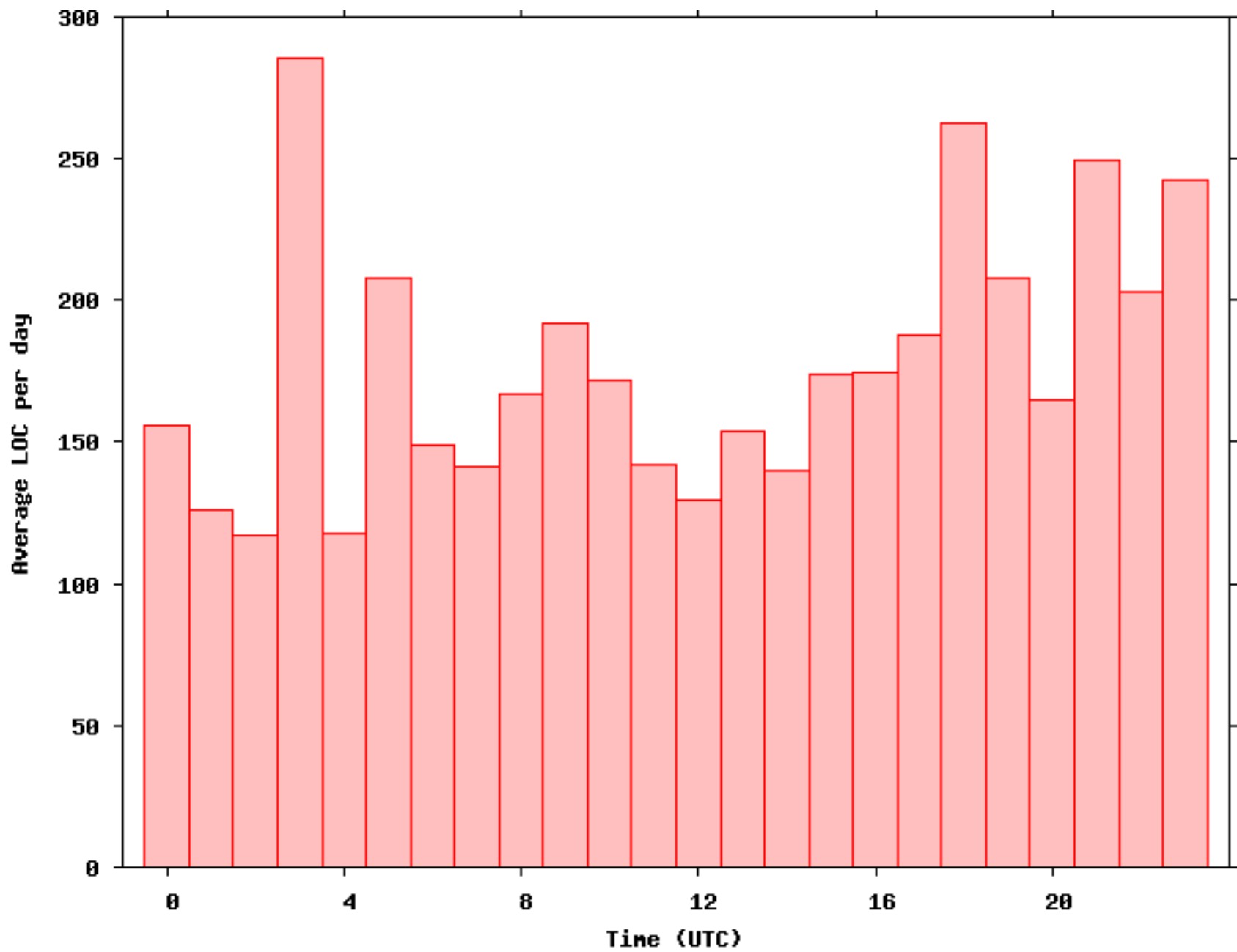
**cvs annotate**



```
$ cvs annotate f |  
awk '{print $2}' |  
sort |  
uniq -c |  
sort -rn
```



1652	spy
1104	polina
827	ajk
372	dds
279	panos
124	mastorad
75	mm
16	nd
2	nmil
1	pappas



**find**

```
find -name foo
```

```
find -type f
```

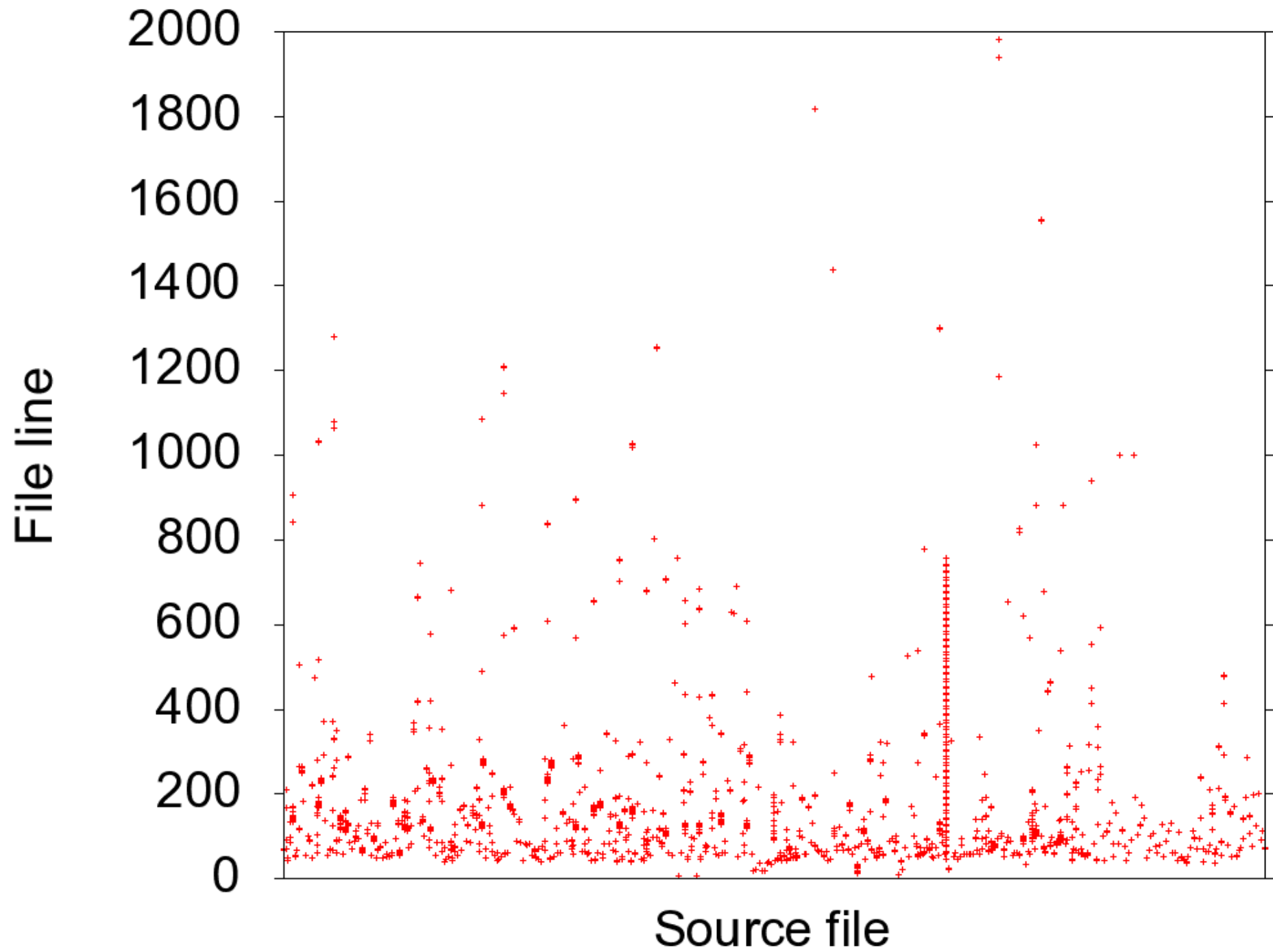
```
find -type d
```

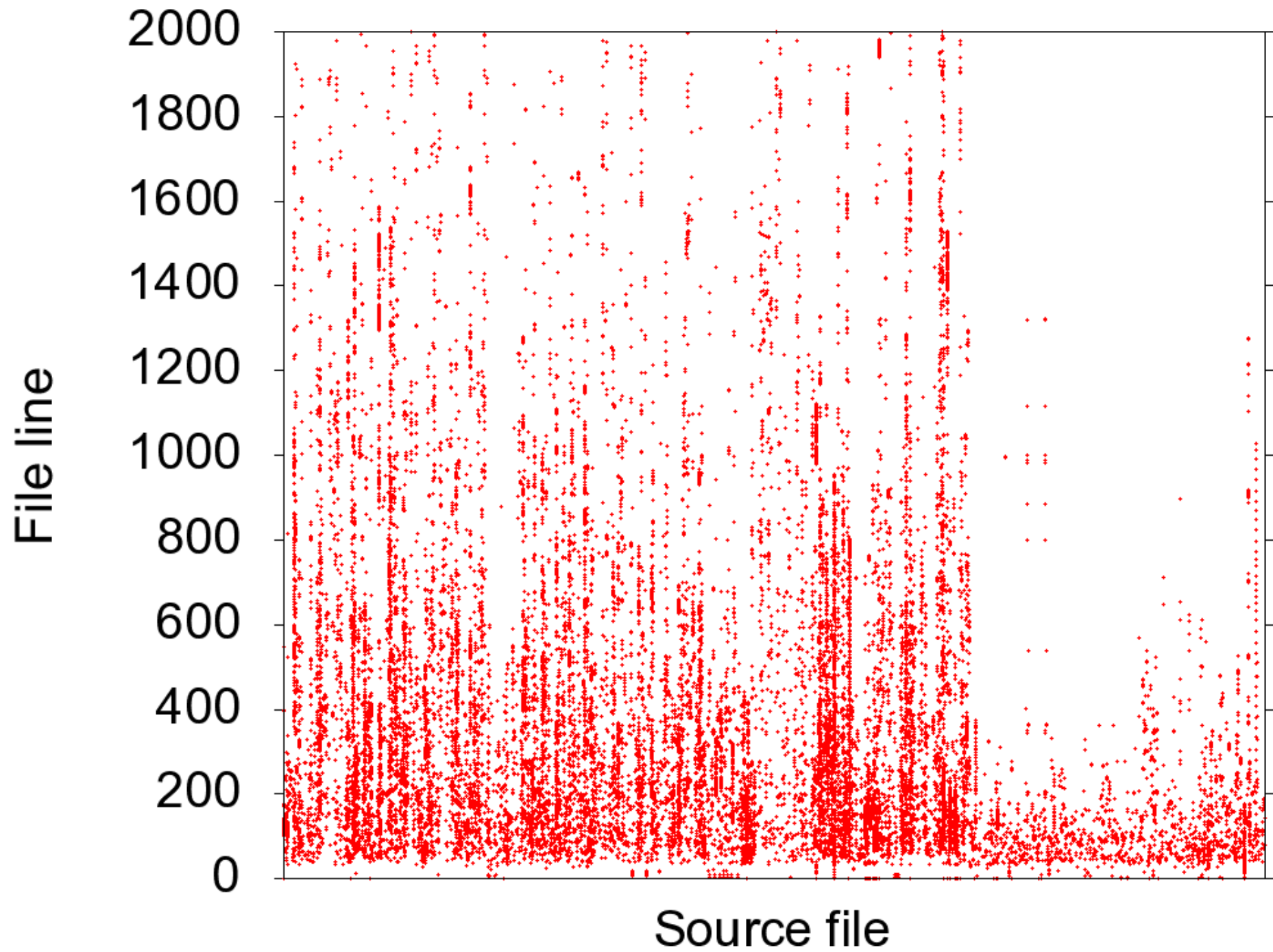
```
find -print0
```

```
$ find build \  
-name '*.class' \  
-type f \  
-mtime -7
```

**nm**

**dumpbin**







```
$ nm -l netbsd |  
awk '/ [BDGRS] / {print $3}' |  
sed s/_//
```

ar

1dd

```
$ cd /usr/bin
$ ldd * 2>/dev/null |
awk '/=>/{print $3}' |
sort |
uniq -c |sort -rn |
head
```

```
372 /lib/libc.so.7
 35 /lib/libncurses.so.7
 33 /lib/libm.so.5
 31 /lib/libz.so.4
 29 /lib/libcrypt.so.4
 27 /lib/libutil.so.7
 27 /lib/libcrypto.so.5
 22 /usr/lib/libstdc++.so.6
 22 /lib/libgcc_s.so.1
 20 /usr/lib/libbz2.so.3
```

541 /lib/libc.so.6

541 /lib/ld-linux.so.2

104 /lib/libdl.so.2

92 /lib/libm.so.6

62 /lib/libncurses.so.5

59 /lib/libnsl.so.1

50 /lib/libcrypt.so.1

34 /usr/lib/libstdc++-

libc6.2-2.so.3

34 /lib/libpam.so.0

28 /usr/lib/libz.so.1

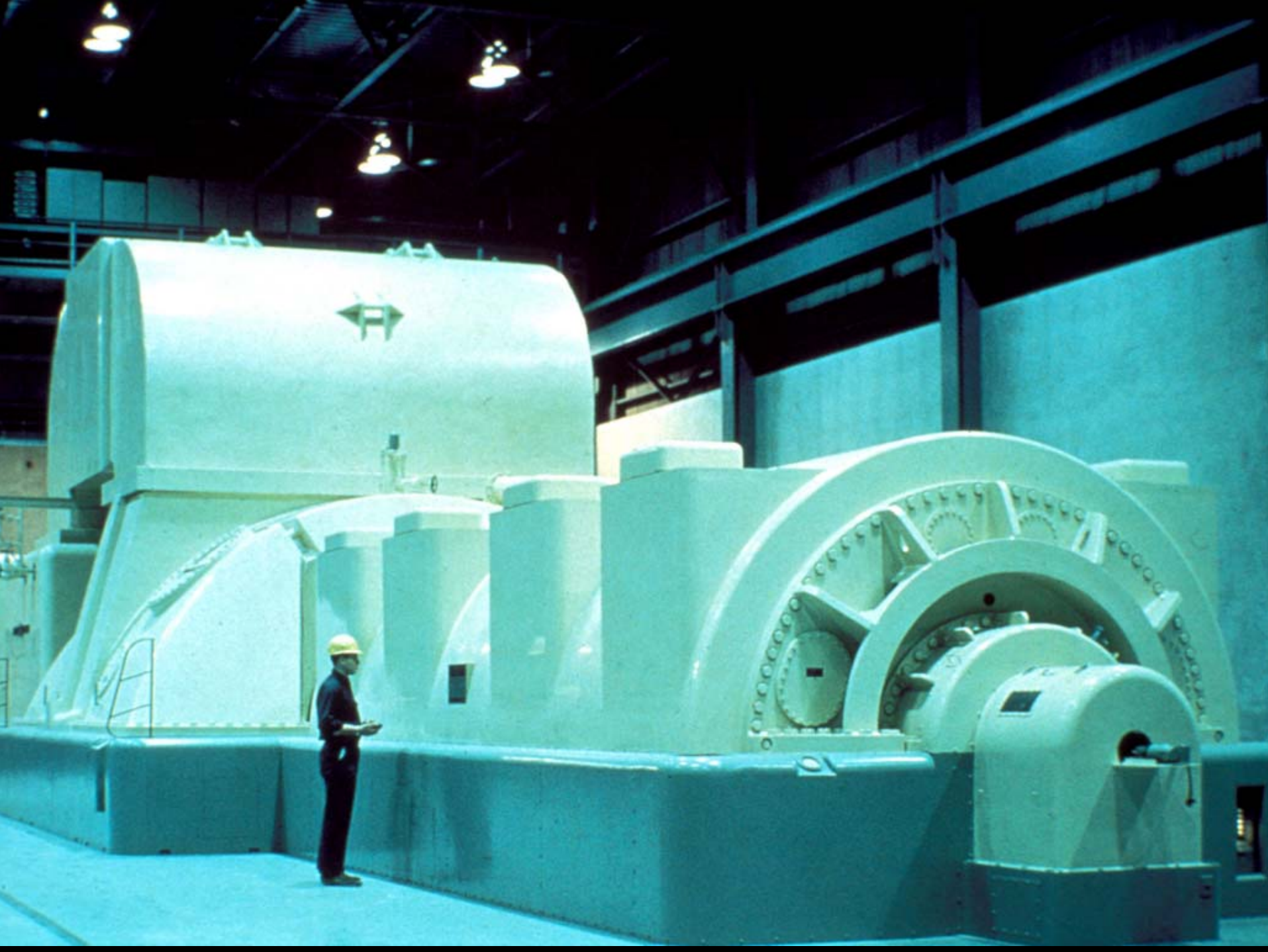
tar

jar

```
$ tar cf - . |  
tar -C dir -xpf -
```

```
$ jar tf ant.jar |  
sed -n \  
' /\.class$/s/^/ant.jar /p' |  
java -jar ckjm.jar
```





jot

\$ jot - 1 101

2

3

4

5

6

7

8

9

10

```
$ grep `jot -s "" -b . 80`
```

dd

```
$ time \  
dd if=/dev/zero \  
bs=1M count=1000 |  
nc -q 0 host 9999
```

**yes**





**grep**

a

.

k\*

[a-z]

[^a-z]

^a

b\$

\. \[ \\*

(a.b) \1

a|b c+ d? {9} {,9}

```
$ egrep '[^aeiouyAEIOUY]{4}' \  
/usr/share/dict/words
```

archchronicler

bergschrund

Eschscholtzia

fruchtschiefer

latchstring

lengthsman

Nachschlag

postphthisic

veldtschoen

Knightsbridge

```
$ egrep \  
'^(.)(.)((.)\4)?\2\1$' \  
/usr/share/dict/words
```

boob

deed

noon

peep

redder

sees

**awk**

```
/^x/ {print $2}
```

```
$2 == $3 {a[$2]++}
```

```
END {print x}
```

```
-F:
```



```
$ tar tvf vmmemctl.tar |  
awk ' {s += $3}  
      END {print s} '
```

**sed**

```
s/if(/if (/;s/do{/do {/  
/^ \*/s/ *\*$//
```

```
/SCCS/d
```

```
-n
```

```
s/^struct \(.*\) {\1/p
```

```
$ ls |  
sed -n \  
's/\(Beetles - One - \)\(.*\)/\  
mv "\1\2" "\2"/p' |  
sh
```

cut

**xm1**

**xgawk**

```
<xsl:param name="today" />
```

```
$ xml tr report.xslt \  
-s today=`date +%Y0101` \  
file.xml >brochure.html
```





sort

$-n$

$-r$

$+2nr$

$-u$

$-t :$

```
$ tar tvf vmmemctl.tar |  
sort +2n
```

**diff**

-b

-u

```
$ diff file1 file2 |  
grep '^ [<>]' |  
wc -l
```

**Comm**

```
$ comm Linux FreeBSD
```

```
[
```

```
arch
```

```
ash
```

```
bash
```

```
cat
```

```
chflags
```

```
chgrp
```

```
chio
```

```
chmod
```

```
chown
```

```
cp
```

```
cpio
```

```
csch
```

```
date
```

```
dd
```

```
df
```

```
dir
```

```
dmesg
```

```
dnsdomainname
```

```
domainname
```

```
echo
```

```
ed
```

```
egrep
```

```
expr
```

```
false
```

```
getfacl
```

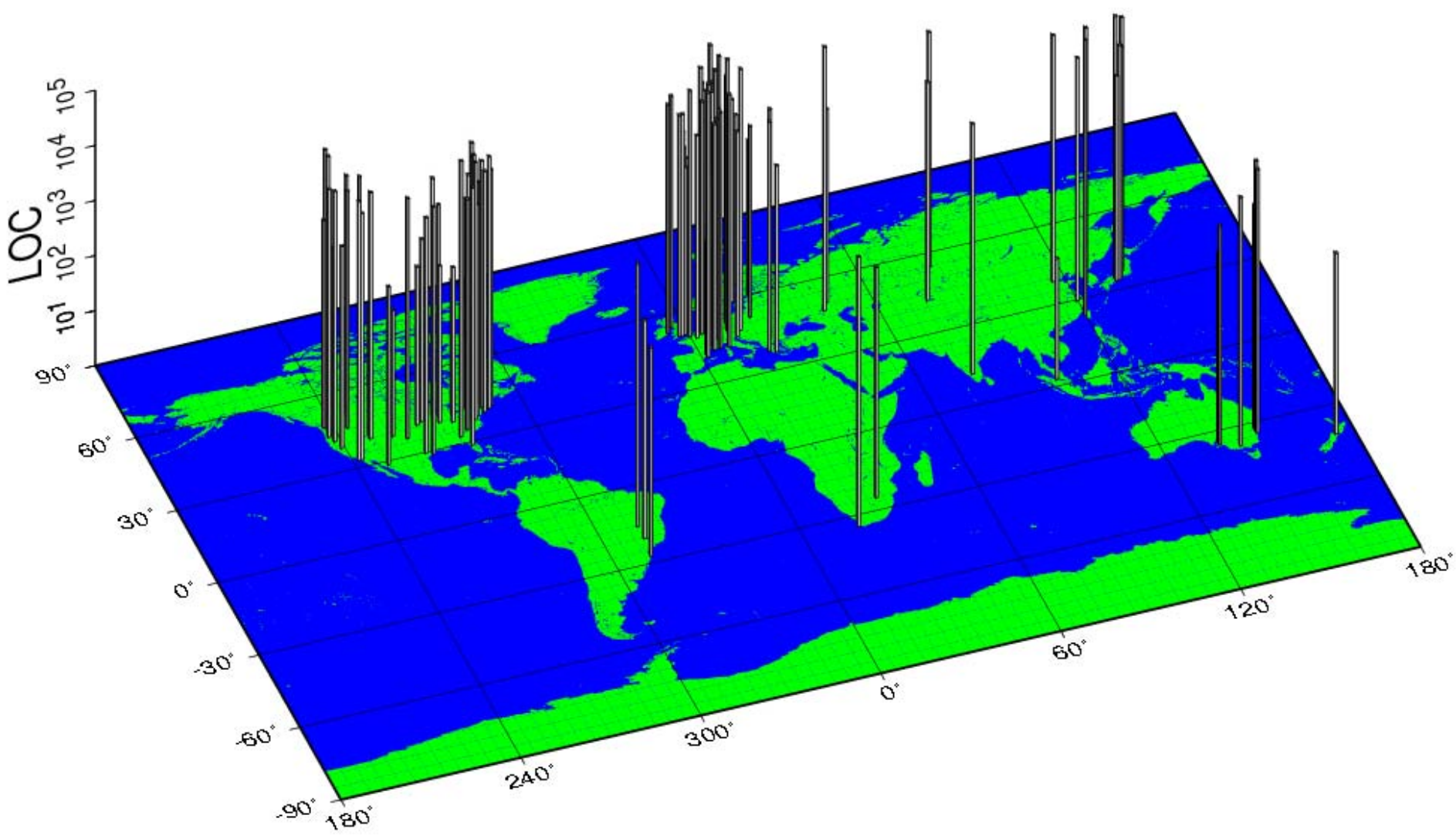


- 1

- 2

- 3

```
$ comm -23 active-committers  
developers-known \  
>developers-unknown
```



rs

\$ jot -r 100 |

rs 10 10

65	100	84	36	4	98	18	64	47	15
46	54	35	50	49	23	37	41	79	31
26	71	15	34	74	63	43	18	49	36
35	19	19	11	25	98	35	36	54	75
12	89	72	94	10	6	10	51	37	7
85	43	99	5	86	57	73	24	61	95
37	84	41	35	38	93	60	30	84	39
94	44	49	35	27	67	37	85	42	96
77	14	43	6	92	57	88	18	31	20
86	100	97	8	70	53	24	35	36	17

```
$ rs -ec 0 4 |  
sort |  
rs -c 0 1
```

tr

```
$ tr -C '[a-zA-Z]' '\n' </etc/motd |  
tr A-Z a-z |  
sort |  
uniq |  
grep -v '[sd]$\n' |  
comm -23 - /usr/share/dict/words
```



**recode**

tac

rev

join

date



A VICIOUS RACKET  
WITH ITS ARMS  
AROUND YOUR  
CHILDREN!

DEBAUCHERY

THE TRUTH  
ABOUT . . .

**MARIJUANA**

THE SMOKE OF HELL!

“**DEVIL'S  
HARVEST**”

```
date -f'%b %d %Y' -j +%s' $D || # Mar 12 2004
date -f'%d %B %Y' -j +%s' $D || # 12 March 2004
date -f'%a %d %B %Y' -j +%s' $D || # Mon 12 March 2004
date -f'%+' -j +%s' $D || # Thu May 8 00:14:31 PDT 1997
date -f'%Y-%m-%d' -j +%s' $D || # 2002-05-17
date -f'%Y.%m.%d' -j +%s' $D || # 2001.02.08
date -f'%Y/%m/%d' -j +%s' $D || # 2001/07/01
date -f'%y/%m/%d' -j +%s' $D || # 01/09/18
date -f'%d-%b-%Y' -j +%s' $D || # 25-Sep-2001
date -f'%d.%m.%Y' -j +%s' $D || # 03.02.2002
date -f'%m-%d-%Y' -j +%s' $D || # 03-02-2002
date -f'%m/%d/%Y' -j +%s' $D || # 2/22/1999
```

gvpr



```
BEG_G {  
    $tvtype = TV_fwd;  
    $tvroot = node($, ARGV[0]);  
}  
N [$tvroot = NULL; 1]  
END_G {  
    induce($T);  
    write($T);  
    exit(0);  
}
```



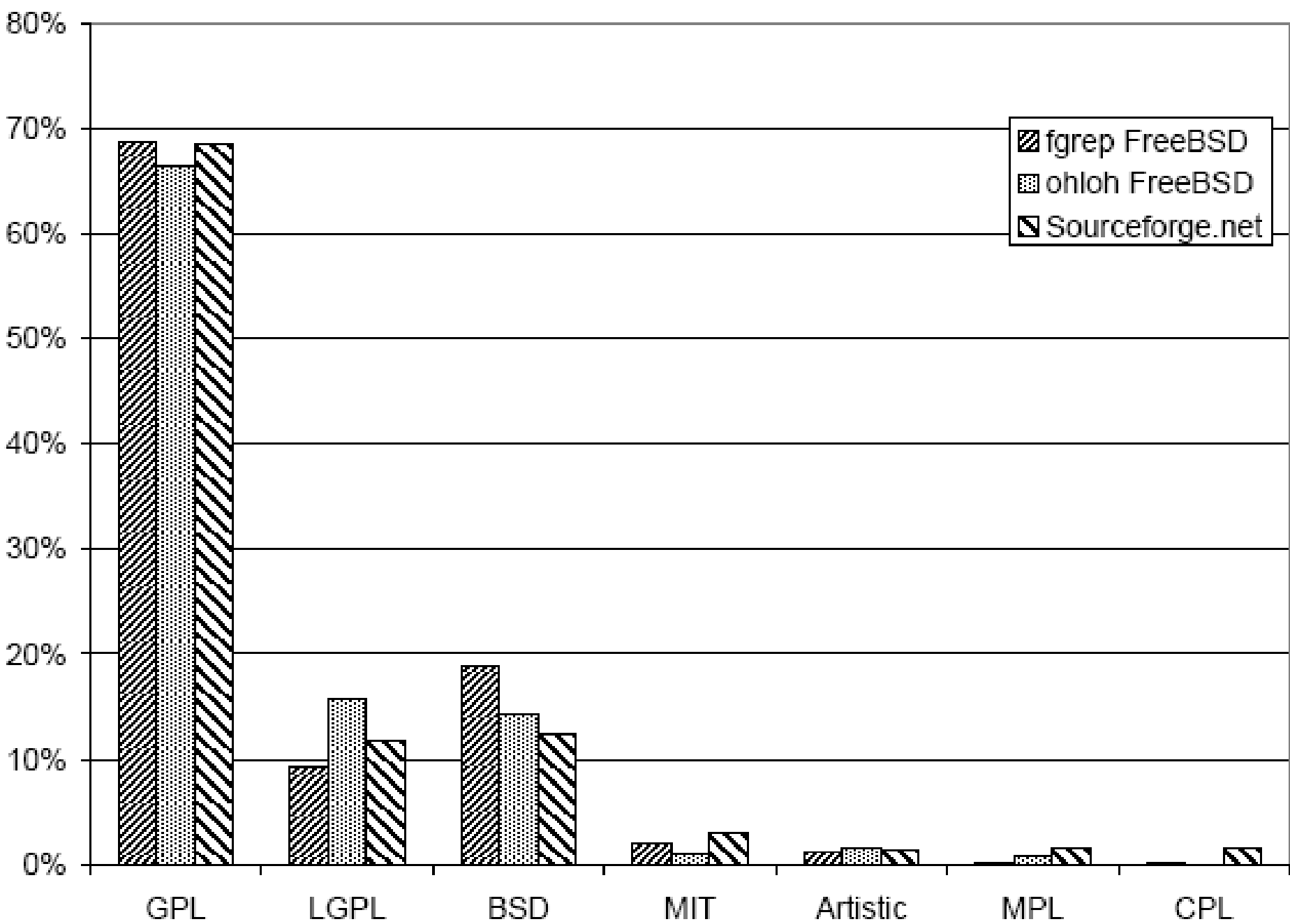
**WC**

**head**

tail

uniq

```
awk '{print $2}' licenses |  
sort |  
uniq -c |  
sort -rn
```

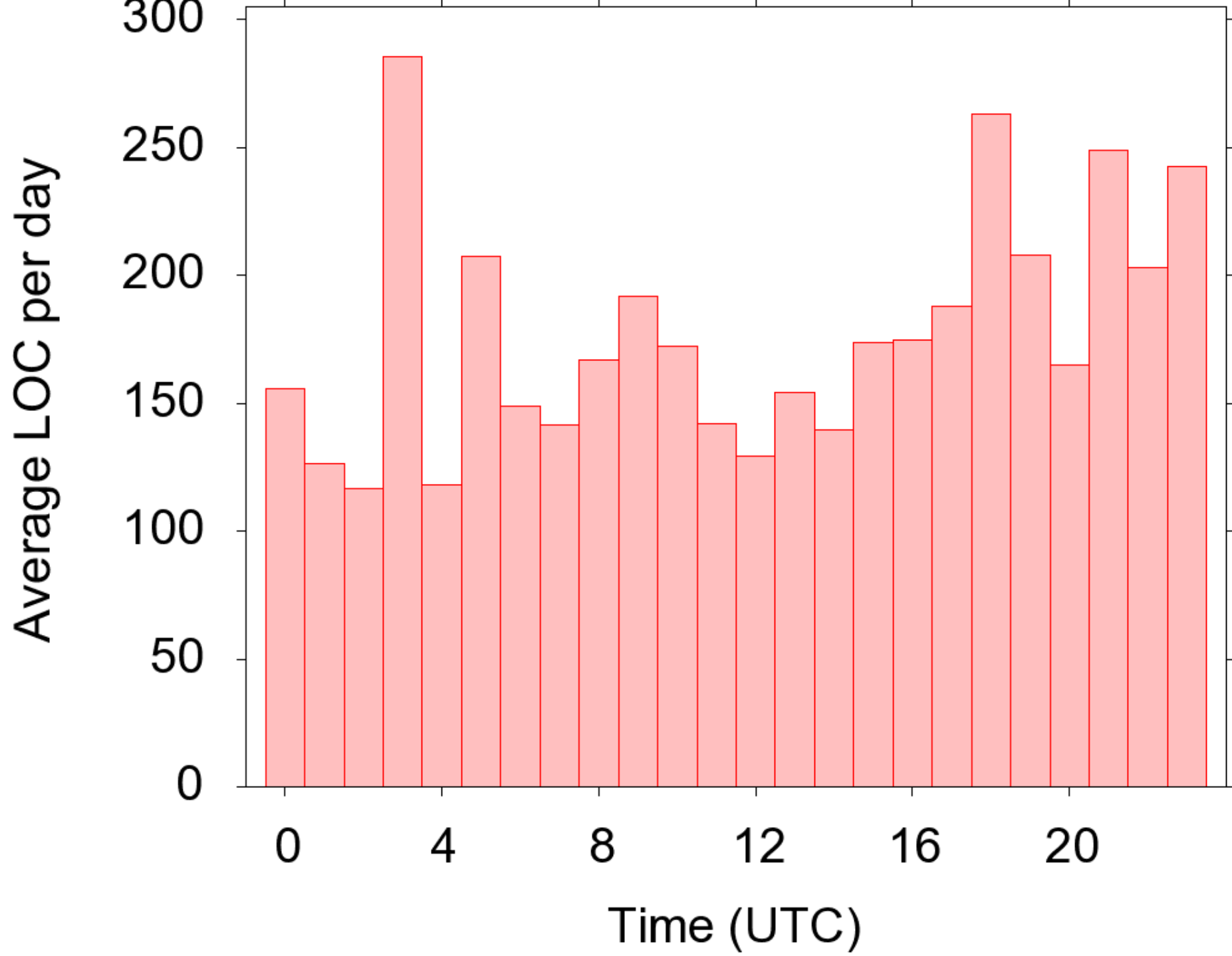




**fmt**

**awk**

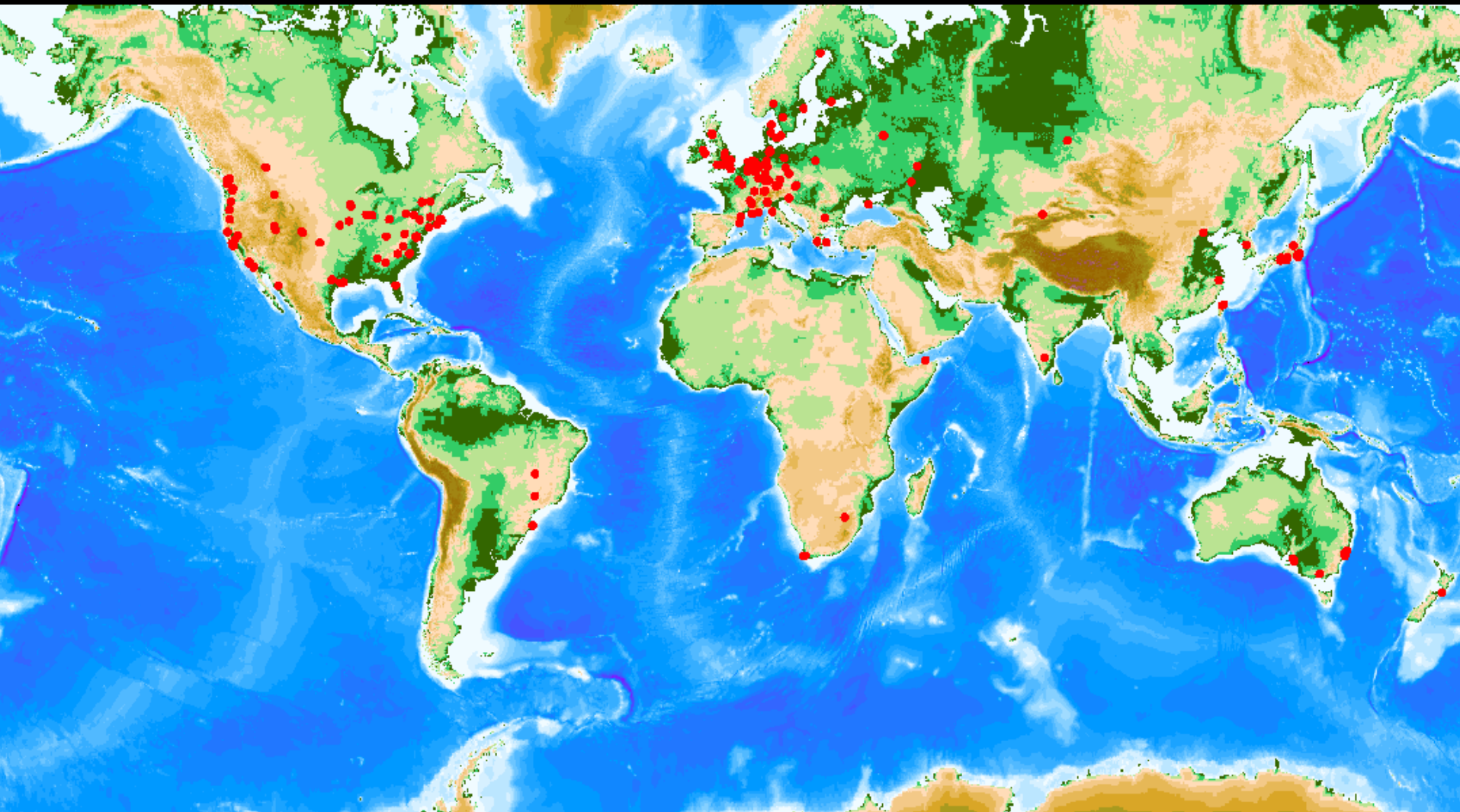
```
find src -type f |
grep -v CVS |
xargs cvs log -SN |
sed -n '/^date:/s/[+;]//gp' |
awk '{hlines[$5] += $9}
     END {for (i in hlines)
           print i, hlines[i}}' |
sort >hlines
```



**gnuplot**

```
set xlabel 'Time (UTC)'  
set ylabel 'Average LOC per day'  
set style fill solid 0.250000  
border  
set tics out  
set xtics 0,4,20  
plot [-1:24] [0:305]  
'roundclock.dat' notitle with  
boxes
```

**GMT**

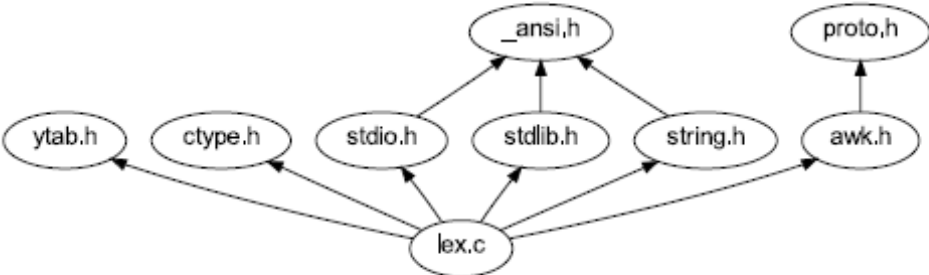




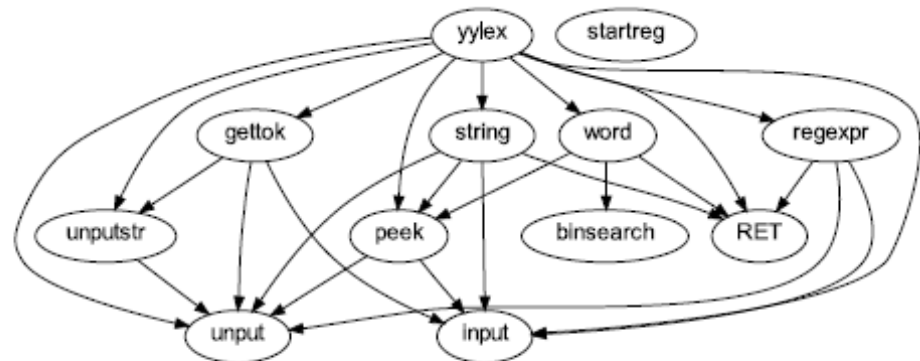
dot

```
digraph gvpr {  
  "atk" -> "gettext";  
  "gettext" -> "libiconv";  
  "libxml2" -> "libiconv";  
  "gtk20" -> "gettext";  
  "firefox" -> "glib20";  
  "firefox" -> "atk";  
}
```

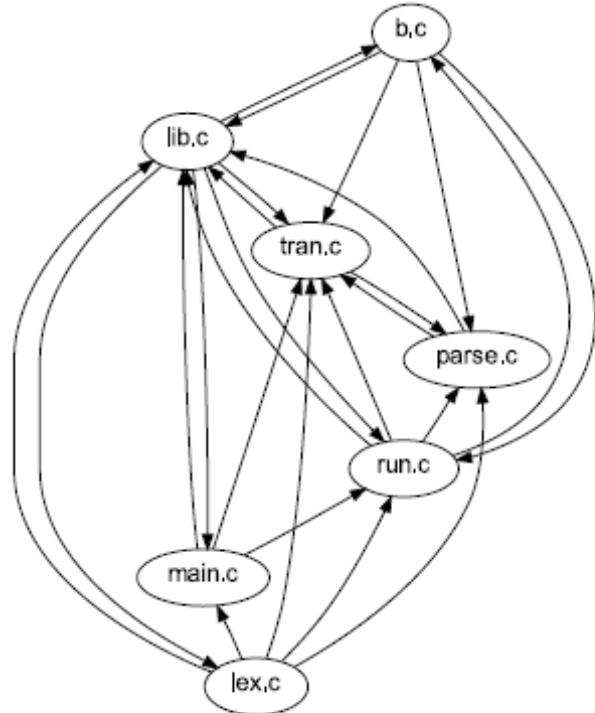




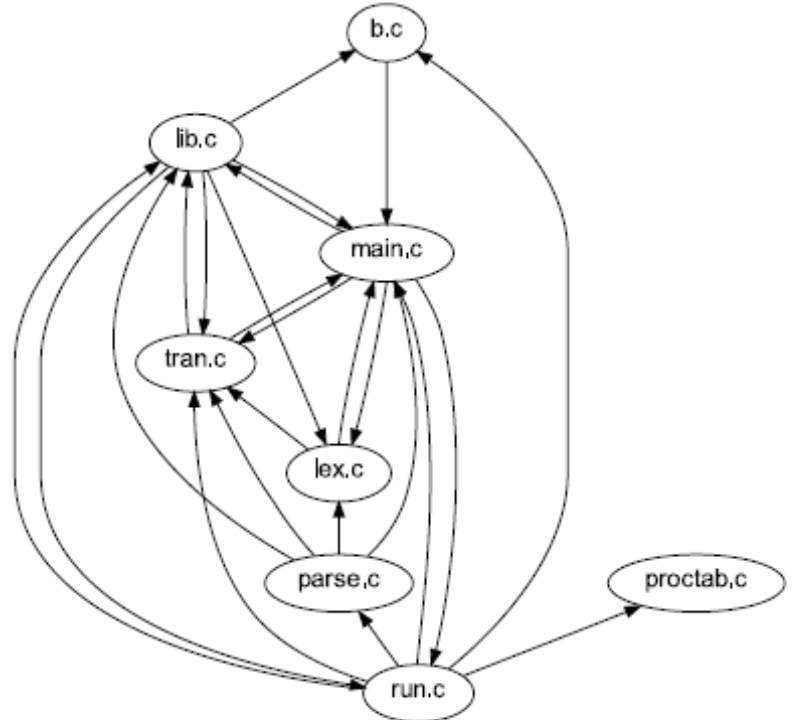
(a) Included files.



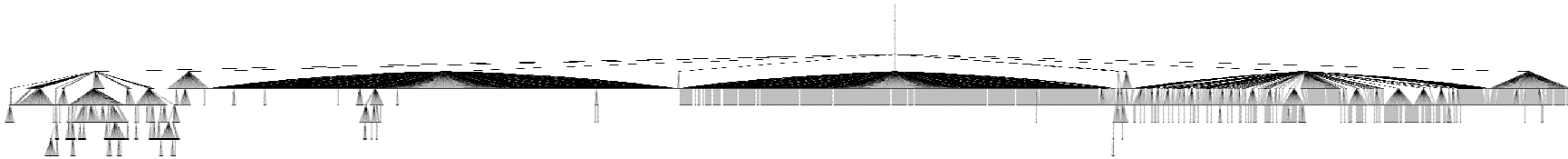
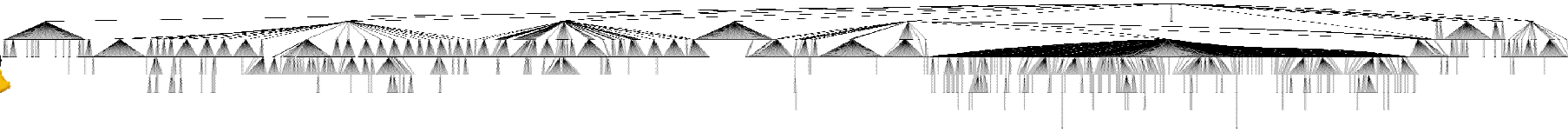
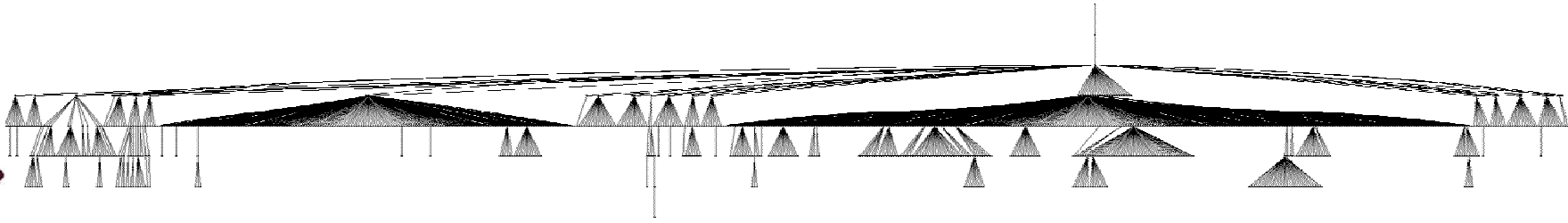
(b) Call graph spanning functions and macros.

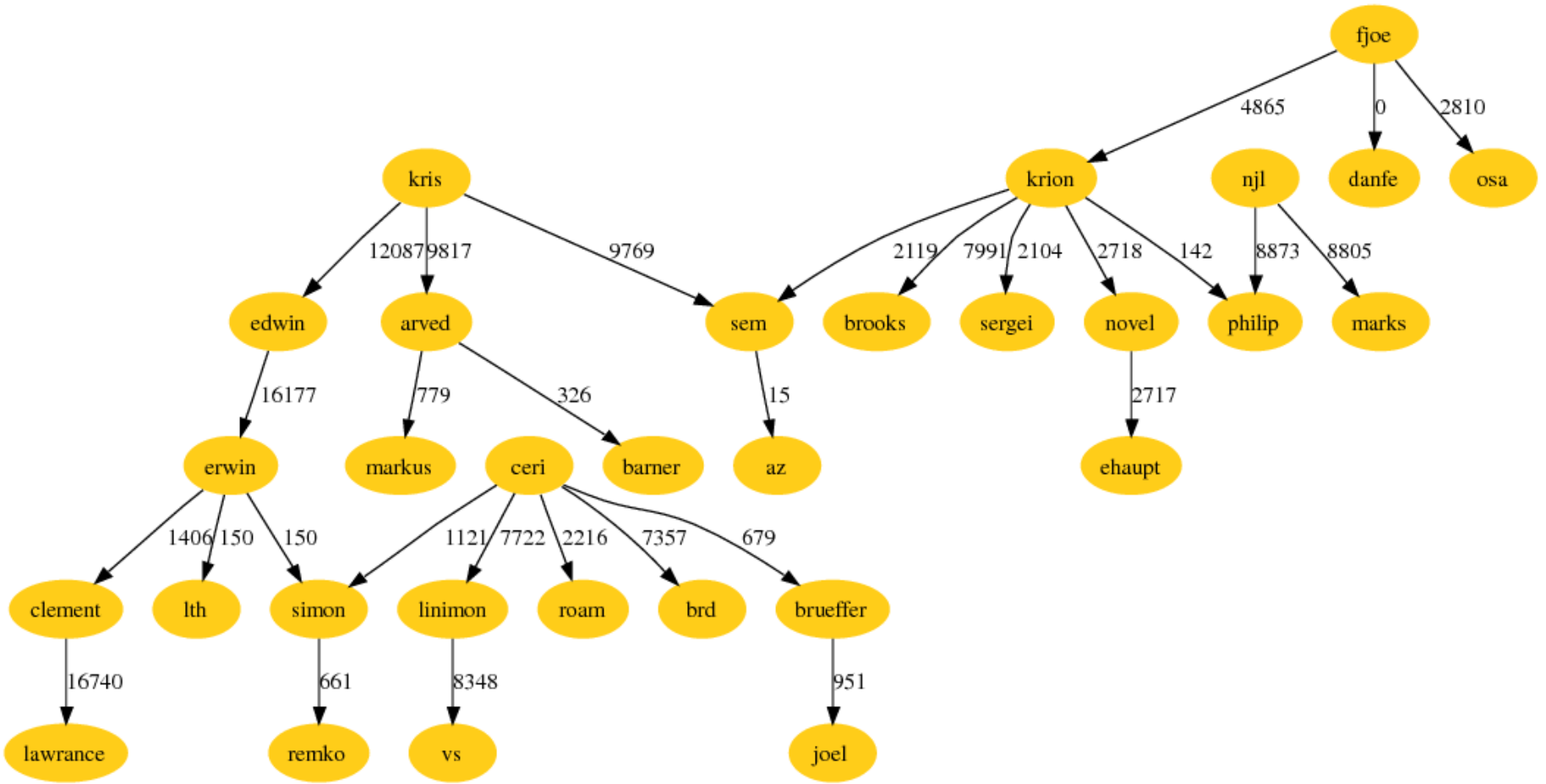


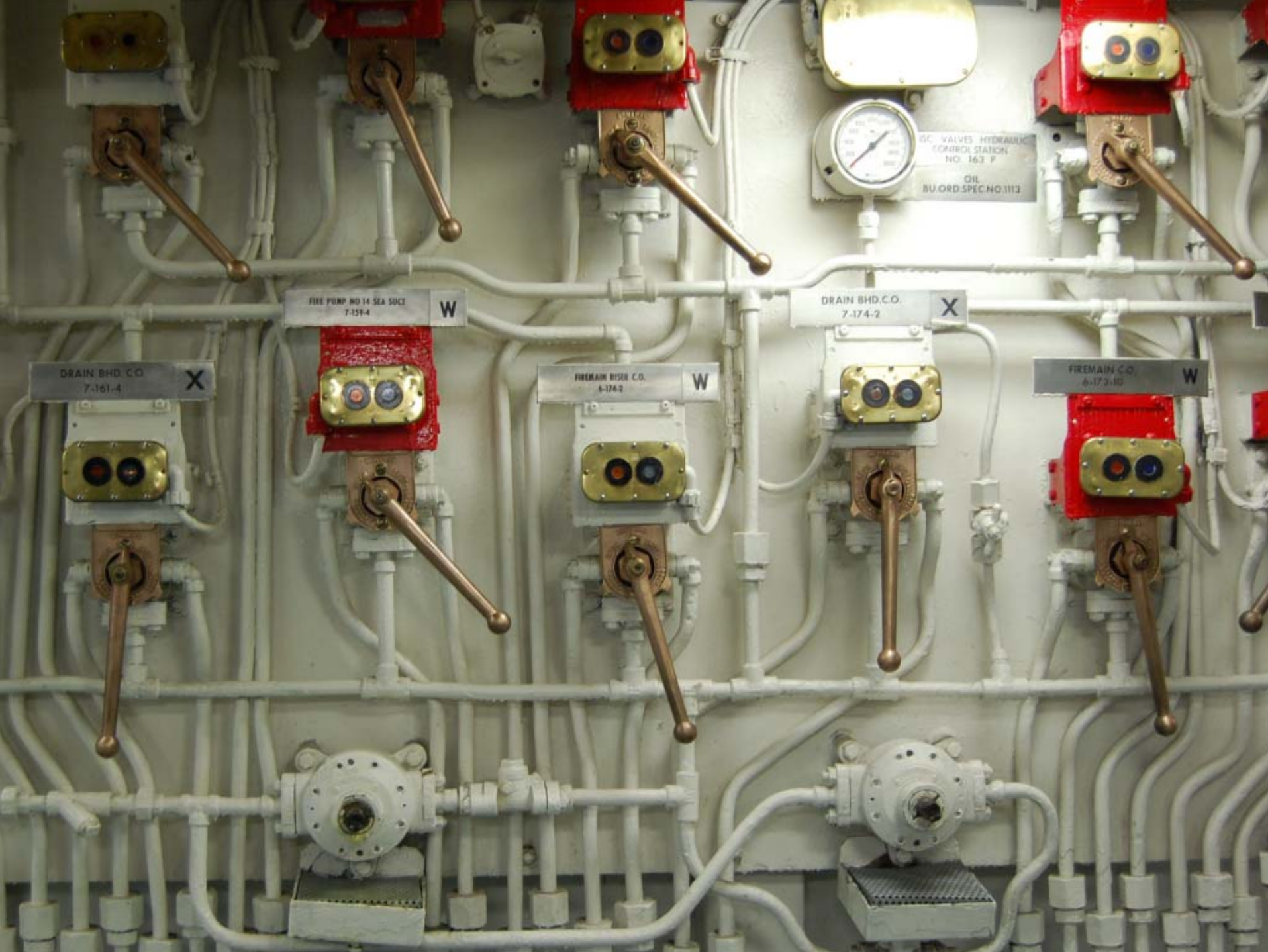
(c) Control dependencies between files.



(d) Data dependencies between files.







H.C. VALVES HYDRAULIC  
CONTROL STATION  
NO. 163 P  
OIL  
BU ORD SPEC NO 1113

FIRE PUMP NO 14 SEA SUCT  
7-159-4 W

DRAIN BHD.CO.  
7-174-2 X

DRAIN BHD.CO.  
7-161-4 X

FIREMAN RISER C.O.  
6-174-2 W

FIREMAN C.O.  
6-173-10 W





generate |

tee file |

summarize

**xargs**

- 0

- P

- n 1

( ... )

|

...|

**while read x**

```
if cmd
```

```
then
```

```
...
```

```
fi |
```



sh



```
xargs grep RE /dev/null
```

echo **Error 1>&2**

```
if [ x$a = x ]  
then  
    echo '$a was empty'  
fi
```

exec

```
mv file dir/
```

**sed**

**...**

**|**

**sh**

```
TMPDIR="$ {TMP:- /tmp} /$$"  
trap 'rm -rf "$TMPDIR"' 0  
trap 'exit 2' 1 2 15  
mkdir "$TMPDIR"
```

Perl

Python

Ruby

Groovy





cat file |  
command

**command**

**<file**

```
command
```

```
if [ $? -ne 0 ]
```

```
then
```

```
if !command  
then
```

`grep RE |`  
`awk ' {... } '`

**awk ' /RE/ { ... } '**

```
echo `ls` |  
sed `s/ /:/`
```



```
ls |
```

```
paste -sd:
```

```
echo $VAR |  
sed 's/\([^-\]*\).*\/\1/'
```

`expr "$VAR" : '\([^-\]*\)'`

```
find /foo -name ... |  
while read filename  
do  
    ls -ld $filename  
done
```

```
find /foo -name ... \  
-exec ls -ld '{}' \;
```

```
find /foo -name ... |  
xargs ls -ld
```

# www.spinellis.gr

## twitter.com/DSpinellis

### dds@aueb.gr

