A Survey of Linux Measurement and Diagnostic Tools

When the performance of your embedded Linux device is inadequate, how do you understand what the problem is?

This presentation will provide an overview of some of the available tools to measure and analyze the behavior and resource usage of the Linux kernel and userland applications.

Frank Rowand, Sony Corporation of America

October 16, 2009

Please ask questions during the presentation!

If an answer will be long, I might defer it to the end of the talk.

If I think I am running out of time, I might defer all questions to the end of the talk.

I will be available for discussion and questions after the talk.

Performance Instrumentation is An Active Area of Development

Most examples for today created on:

- Linux kernel
  - < 2.6.29
  - 2.6.29.4
  - 2.6.30
  - 2.6.32-rc3
- x86 hardware
  - 32 bit, 64 bit
  - Intel, AMD
  - SMP, UP

Preview

Some past, current and future tools to investigate:

- Waiting for resources
- Resource usage
- Resource allocation

With a bias toward more recent tools.

Suggestions from Google

linux performance monitoring tools

Suggestions from Google

linux performance monitoring tools

```
Conky                       lsof        ps              tcdump
GKrellM                     mpstat      sa              tcpdump
Ksysguard                   mtr         sa2             time
bonnie                      nagios      sadc            top
cacti                       netperf     sal             traceroute
free                        netstat     sar             uptime
gnome system monitor        nfsstat     smartmontools   vmstat
htop                        nmap        smem            w
iostat                      ntop        spray           wireshark
iozone                      oprofile    ss              xload
iptraf                      ping        strace          xosview
isag                        pmap        sysstat         /proc
```

Suggestions from Google

linux performance monitoring tools

Those tools are useful, but mostly not included in this presentation...

LatencyTop [latency]

“Skipping audio, slower servers, everyone knows the symptoms of latency. But to know what's going on in the system, what's causing the latency, how to fix it... that's a hard question without good answers right now.”

source: http://www.latencytop.org/

LatencyTop [latency]

“LatencyTOP focuses on the cases where the applications want to run and execute useful code, but there's some resource that's not currently available (and the kernel then blocks the process).”

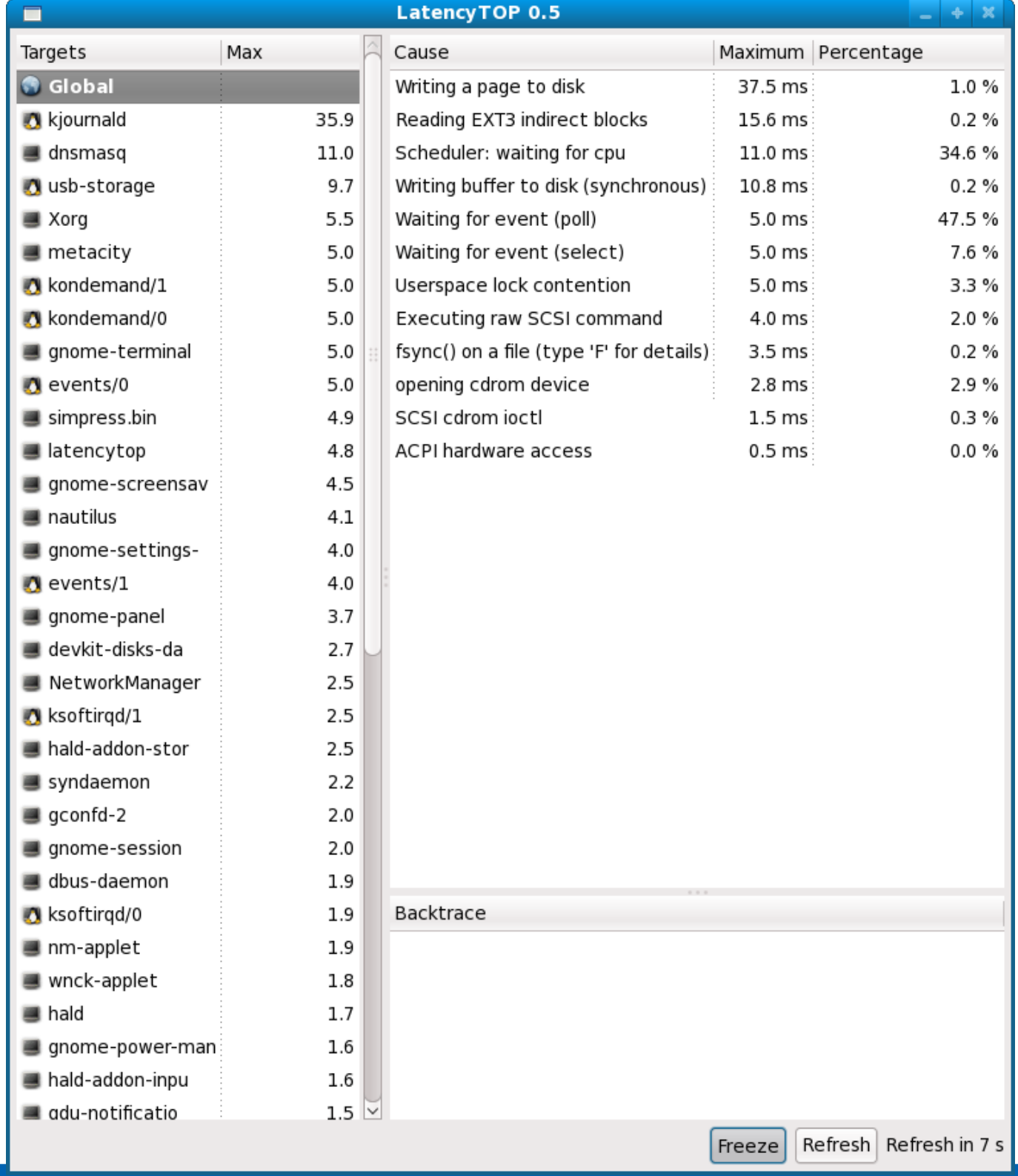
source: http://www.latencytop.org/

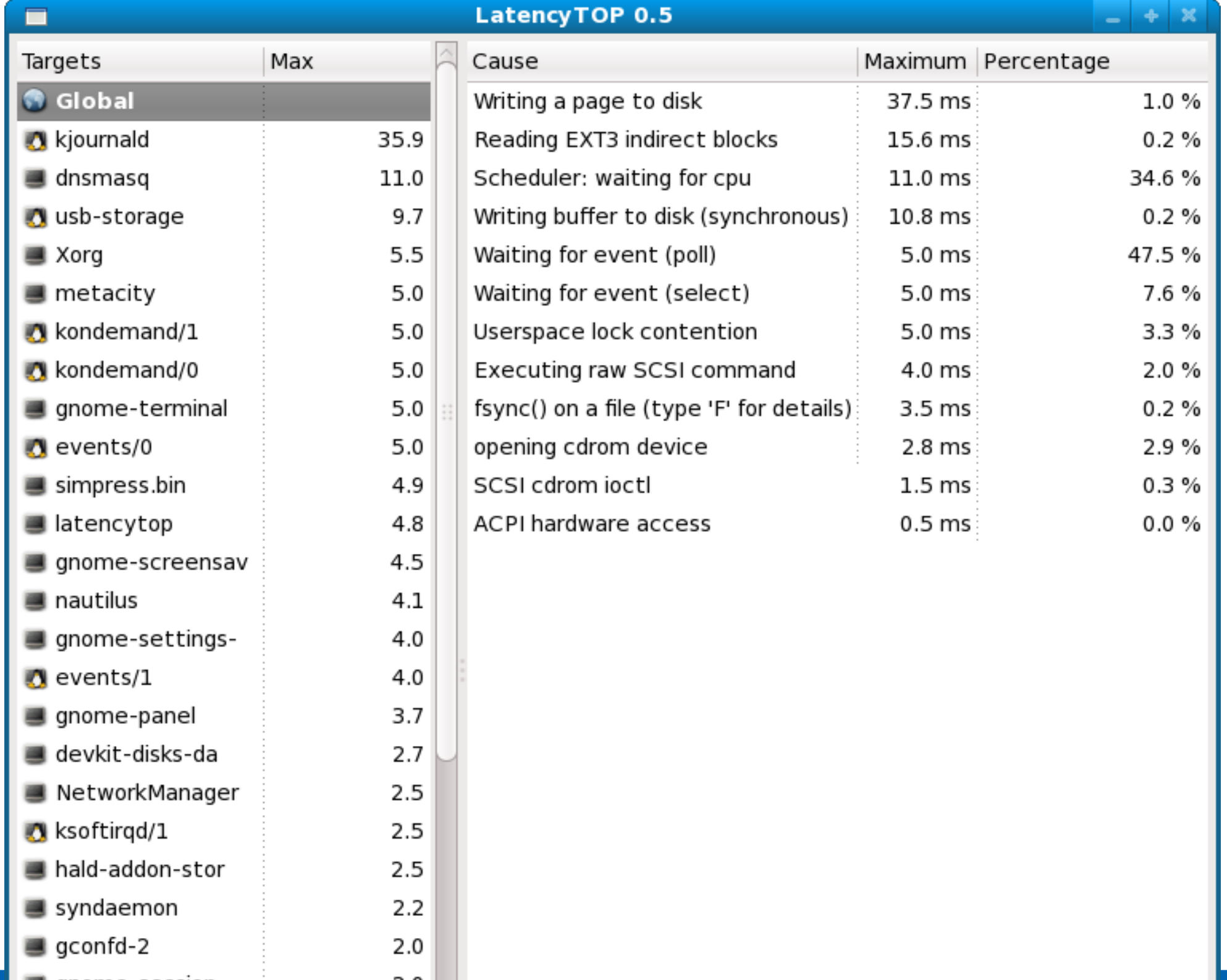
LatencyTop [latency]

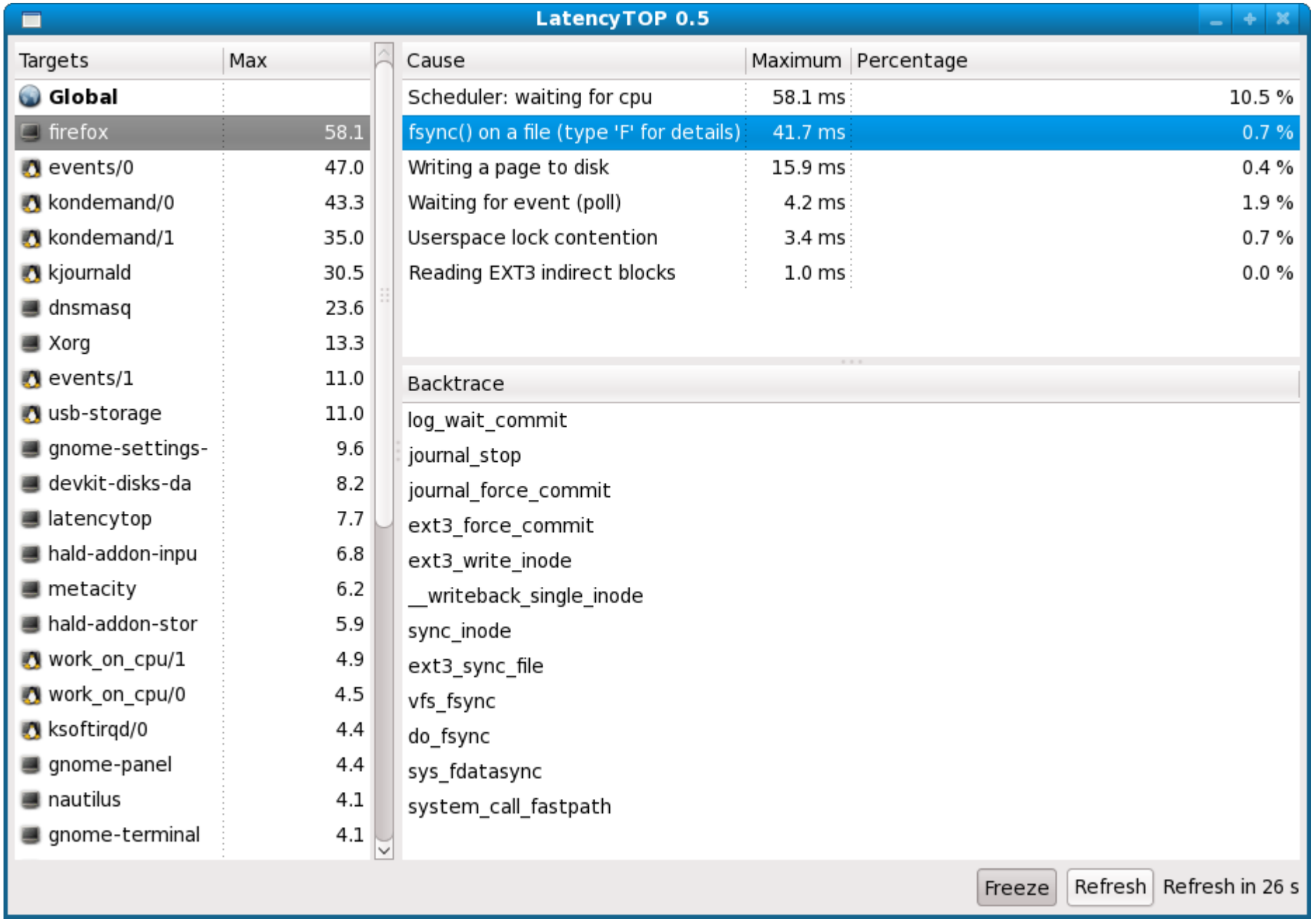
“LatencyTOP focuses on the cases where the applications want to run and execute useful code, but there's some resource that's not currently available (and the kernel then blocks the process).

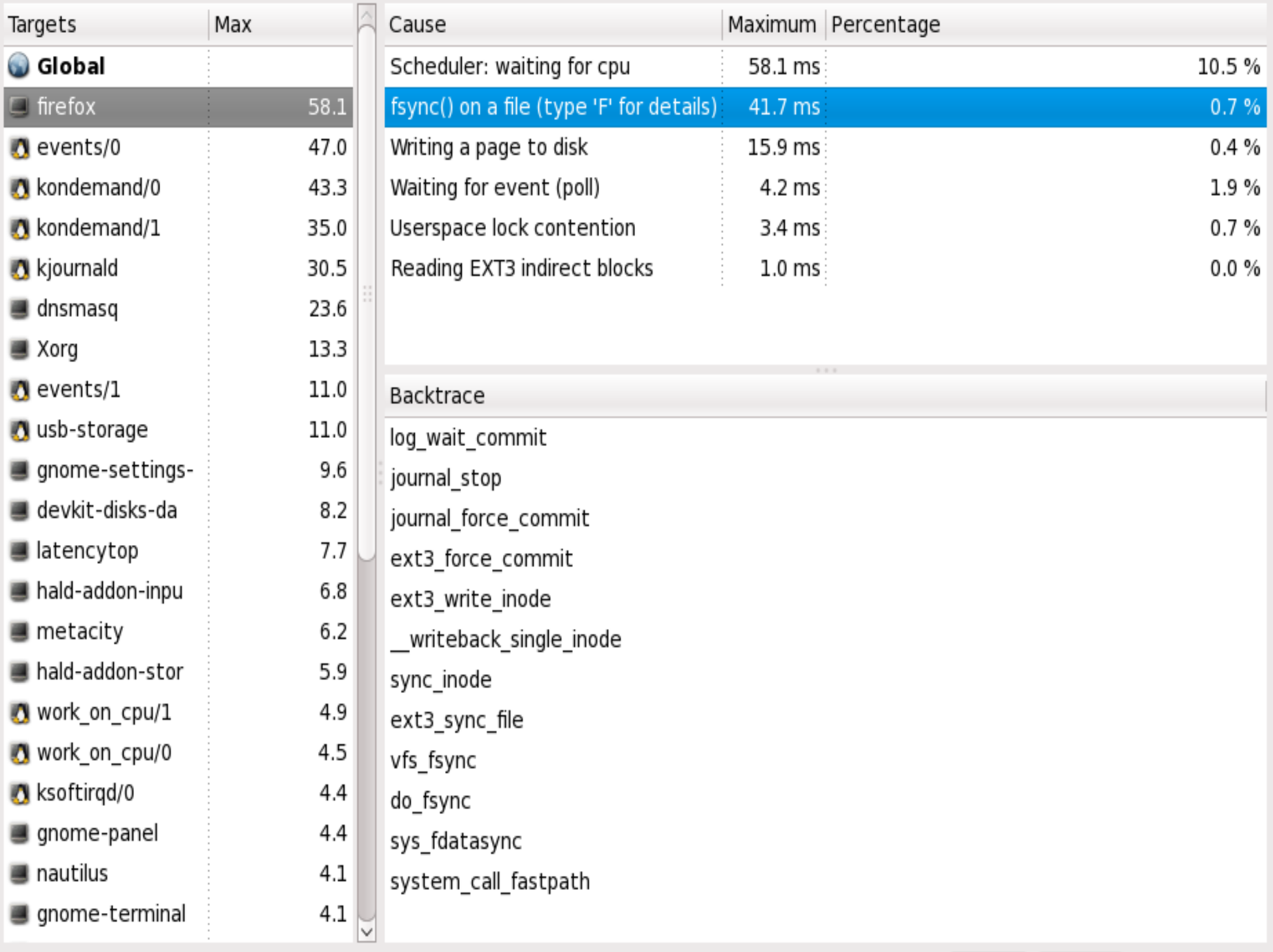
This is done both on a system level and on a per process level, so that you can see what's happening to the system, and which process is suffering and/or causing the delays.”

source: http://www.latencytop.org/


LatencyTOP 0.5
Targets
Max
Global
kjournald 35.9
dnsmasq 11.0
usb-storage 9.7
Xorg 5.5
metacity 5.0
kondemand/1 5.0
kondemand/0 5.0
gnome-terminal 5.0
events/0 5.0
simpress.bin 4.9
latencytop 4.8
gnome-screensav 4.5
nautilus 4.1
gnome-settings- 4.0
events/1 4.0
gnome-panel 3.7
devkit-disks-da 2.7
NetworkManager 2.5
ksoftirqd/1 2.5
hald-addon-stor 2.5
syndaemon 2.2
gconfd-2 2.0
gnome-session 2.0
dbus-daemon 1.9
ksoftirqd/0 1.9
nm-applet 1.9
wnck-applet 1.8
hald 1.7
gnome-power-man 1.6
hald-addon-inpu 1.6
gdu-notificatio 1.5
Cause
Maximum
Percentage
Writing a page to disk 37.5 ms 1.0 %
Reading EXT3 indirect blocks 15.6 ms 0.2 %
Scheduler: waiting for cpu 11.0 ms 34.6 %
Writing buffer to disk (synchronous) 10.8 ms 0.2 %
Waiting for event (poll) 5.0 ms 47.5 %
Waiting for event (select) 5.0 ms 7.6 %
Userspace lock contention 5.0 ms 3.3 %
Executing raw SCSI command 4.0 ms 2.0 %
fsync() on a file (type 'F' for details) 3.5 ms 0.2 %
opening cdrom device 2.8 ms 2.9 %
SCSI cdrom ioctl 1.5 ms 0.3 %
ACPI hardware access 0.5 ms 0.0 %
Backtrace
Freeze
Refresh
Refresh in 7 s


LatencyTOP 0.5
Targets
Max
Global
kjournald 35.9
dnsmasq 11.0
usb-storage 9.7
Xorg 5.5
metacity 5.0
kondemand/1 5.0
kondemand/0 5.0
gnome-terminal 5.0
events/0 5.0
simpress.bin 4.9
latencytop 4.8
gnome-screensav 4.5
nautilus 4.1
gnome-settings- 4.0
events/1 4.0
gnome-panel 3.7
devkit-disks-da 2.7
NetworkManager 2.5
ksoftirqd/1 2.5
hald-addon-stor 2.5
syndaemon 2.2
gconfd-2 2.0
Cause
Maximum
Percentage
Writing a page to disk 37.5 ms 1.0 %
Reading EXT3 indirect blocks 15.6 ms 0.2 %
Scheduler: waiting for cpu 11.0 ms 34.6 %
Writing buffer to disk (synchronous) 10.8 ms 0.2 %
Waiting for event (poll) 5.0 ms 47.5 %
Waiting for event (select) 5.0 ms 7.6 %
Userspace lock contention 5.0 ms 3.3 %
Executing raw SCSI command 4.0 ms 2.0 %
fsync() on a file (type 'F' for details) 3.5 ms 0.2 %
opening cdrom device 2.8 ms 2.9 %
SCSI cdrom ioctl 1.5 ms 0.3 %
ACPI hardware access 0.5 ms 0.0 %


LatencyTOP 0.5
Targets
Max
Global
firefox 58.1
events/0 47.0
kondemand/0 43.3
kondemand/1 35.0
kjournald 30.5
dnsmasq 23.6
Xorg 13.3
events/1 11.0
usb-storage 11.0
gnome-settings- 9.6
devkit-disks-da 8.2
latencytop 7.7
hald-addon-inpu 6.8
metacity 6.2
hald-addon-stor 5.9
work_on_cpu/1 4.9
work_on_cpu/0 4.5
ksoftirqd/0 4.4
gnome-panel 4.4
nautilus 4.1
gnome-terminal 4.1
Cause
Maximum
Percentage
Scheduler: waiting for cpu 58.1 ms 10.5 %
fsync() on a file (type 'F' for details) 41.7 ms 0.7 %
Writing a page to disk 15.9 ms 0.4 %
Waiting for event (poll) 4.2 ms 1.9 %
Userspace lock contention 3.4 ms 0.7 %
Reading EXT3 indirect blocks 1.0 ms 0.0 %
Backtrace
log_wait_commit
journal_stop
journal_force_commit
ext3_force_commit
ext3_write_inode
__writeback_single_inode
sync_inode
ext3_sync_file
vfs_fsync
do_fsync
sys_fdatasync
system_call_fastpath
Freeze
Refresh
Refresh in 26 s


Targets
Max
Global
firefox 58.1
events/0 47.0
kondemand/0 43.3
kondemand/1 35.0
kjournald 30.5
dnsmasq 23.6
Xorg 13.3
events/1 11.0
usb-storage 11.0
gnome-settings- 9.6
devkit-disks-da 8.2
latencytop 7.7
hald-addon-inpu 6.8
metacity 6.2
hald-addon-stor 5.9
work_on_cpu/1 4.9
work_on_cpu/0 4.5
ksoftirqd/0 4.4
gnome-panel 4.4
nautilus 4.1
gnome-terminal 4.1
Cause
Maximum
Percentage
Scheduler: waiting for cpu 58.1 ms 10.5 %
fsync() on a file (type 'F' for details) 41.7 ms 0.7 %
Writing a page to disk 15.9 ms 0.4 %
Waiting for event (poll) 4.2 ms 1.9 %
Userspace lock contention 3.4 ms 0.7 %
Reading EXT3 indirect blocks 1.0 ms 0.0 %
Backtrace
log_wait_commit
journal_stop
journal_force_commit
ext3_force_commit
ext3_write_inode
__writeback_single_inode
sync_inode
ext3_sync_file
vfs_fsync
do_fsync
sys_fdatasync
system_call_fastpath

LatencyTop [latency]

Version 0.1 announced 18 January 2008

mutrace

[lock]

Mutex lock contention in application programs

mutrace

[lock]

Example: http://0pointer.de/blog/projects/mutrace.html

$ LD_PRELOAD=/home/lennart/projects/mutrace/libmutrace.so gedit
mutrace: 0.1 sucessfully initialized.

```
mutrace: 10 most contended mutexes:

 Mutex #   Locked  Changed    Cont. tot.Time[ms] avg.Time[ms] max.Time[ms]       Type
      35   368268      407      275      120,822        0,000        0,894     normal
       5   234645      100       21       86,855        0,000        0,494     normal
      26   177324       47        4       98,610        0,001        0,150     normal
      19    55758       53        2       23,931        0,000        0,092     normal
      53      106       73        1        0,769        0,007        0,160     normal
      25    15156       70        1        6,633        0,000        0,019     normal
       4      973       10        1        4,376        0,004        0,174     normal
      75       68       62        0        0,038        0,001        0,004     normal
       9     1663       52        0        1,068        0,001        0,412     normal
       3   136553       41        0       61,408        0,000        0,281     normal
     ...      ...      ...      ...          ...          ...          ...        ...

mutrace: Total runtime 9678,142 ms.
```

mutrace

[lock]

```
mutrace: 10 most contended mutexes:

 Mutex #   Locked  Changed    Cont. tot.Time[ms] avg.Time[ms] max.Time[ms]
      35   368268      407      275      120,822        0,000        0,894
       5   234645      100       21       86,855        0,000        0,494
      26   177324       47        4       98,610        0,001        0,150
      19    55758       53        2       23,931        0,000        0,092
      53      106       73        1        0,769        0,007        0,160
      25    15156       70        1        6,633        0,000        0,019
       4      973       10        1        4,376        0,004        0,174
      75       68       62        0        0,038        0,001        0,004
       9     1663       52        0        1,068        0,001        0,412
       3   136553       41        0       61,408        0,000        0,281
     ...      ...      ...      ...          ...          ...          ...

mutrace: Total runtime 9678,142 ms.
```

mutrace [lock]

Example: http://0pointer.de/blog/projects/mutrace.html

Locked - count: calls to lock mutex
Changed - count: owner of the mutex changed
Cont. - count: requestor waited on lock
tot.Time - total time lock held
avg.Time - average time lock held
max.Time - maximum time lock held
Type - recursive, normal, or otherwise

(Cont. == contended)

mutrace

[lock]

- User space mutex only (“Adding support for rwlocks should be easy”)

- “cannot profile mutexes that are used internally in glibc, such as those used for synchronizing stdio and suchlike”

- requires recent Linux distribution

smem [memory]

Reports memory usage

RSS (Resident Set Size) typically overestimates process memory usage because shared memory is double counted for each user.

PSS (Proportional Set Size) apportions shared memory across all of its users.

PSS is the special smem feature.

```
? smem -r -u
User     Count     Swap      USS      PSS      RSS
frowand     55        0   503576   522509   727732
root        54        0   244308   250594   435132
smmsp        1        0     1196     1316     1824
gdm          1        0      220      226      560
```

```
? smem -r -P vim
  PID User     Command                         Swap      USS      PSS      RSS
11039 frowand  /usr/bin/python ./smem -r -         0     3772     4289     5548
 7728 frowand  vim outline                        0     1372     1682     4008
 2603 frowand  vim notes                          0     1244     1550     3868
 7586 frowand  vim outline_bare                   0     1216     1520     3836
 7970 frowand  vim resource_list                  0     1212     1508     3792
 5509 frowand  vim notes_articles                 0     1180     1445     3628
 4748 frowand  vim info_creation_tools            0     1156     1421     3608
```

```
? smem -r -m -P vim
Map                                             PIDs    AVGPSS      PSS
[heap]                                             7       990     6936
<anonymous>                                        7       278     1948
/usr/bin/vim                                       6       278     1669
/usr/lib64/libpython2.6.so.1.0                     7       211     1483
/usr/lib64/perl5/5.10.0/x86_64-linux-thr           6        75      450
[stack]                                            7        34      240
/lib64/libc-2.10.1.so                              7        17      119
/lib64/ld-2.10.1.so                                7         9       63
/lib64/libpthread-2.10.1.so                        7         8       57
/lib64/libdl-2.10.1.so                             7         8       56
/lib64/libselinux.so.1                             6         9       54
/lib64/libnsl-2.10.1.so                            6         9       54
/lib64/libutil-2.10.1.so                           7         7       52
/lib64/libm-2.10.1.so                              7         7       52
/lib64/libresolv-2.10.1.so                         6         8       48
/lib64/libnss_files-2.10.1.so                      6         8       48
/lib64/libncurses.so.5.7                           6         8       48
/lib64/libcrypt-2.10.1.so                          6         8       48
/lib64/libtinfo.so.5.7                             6         7       42
/lib64/libacl.so.1.1.0                             6         6       38
/usr/lib64/libgpm.so.2.1.0                         6         5       30
/lib64/libfreebl3.so                               6         5       30
/lib64/libattr.so.1.1.0                            6         4       24

......
```

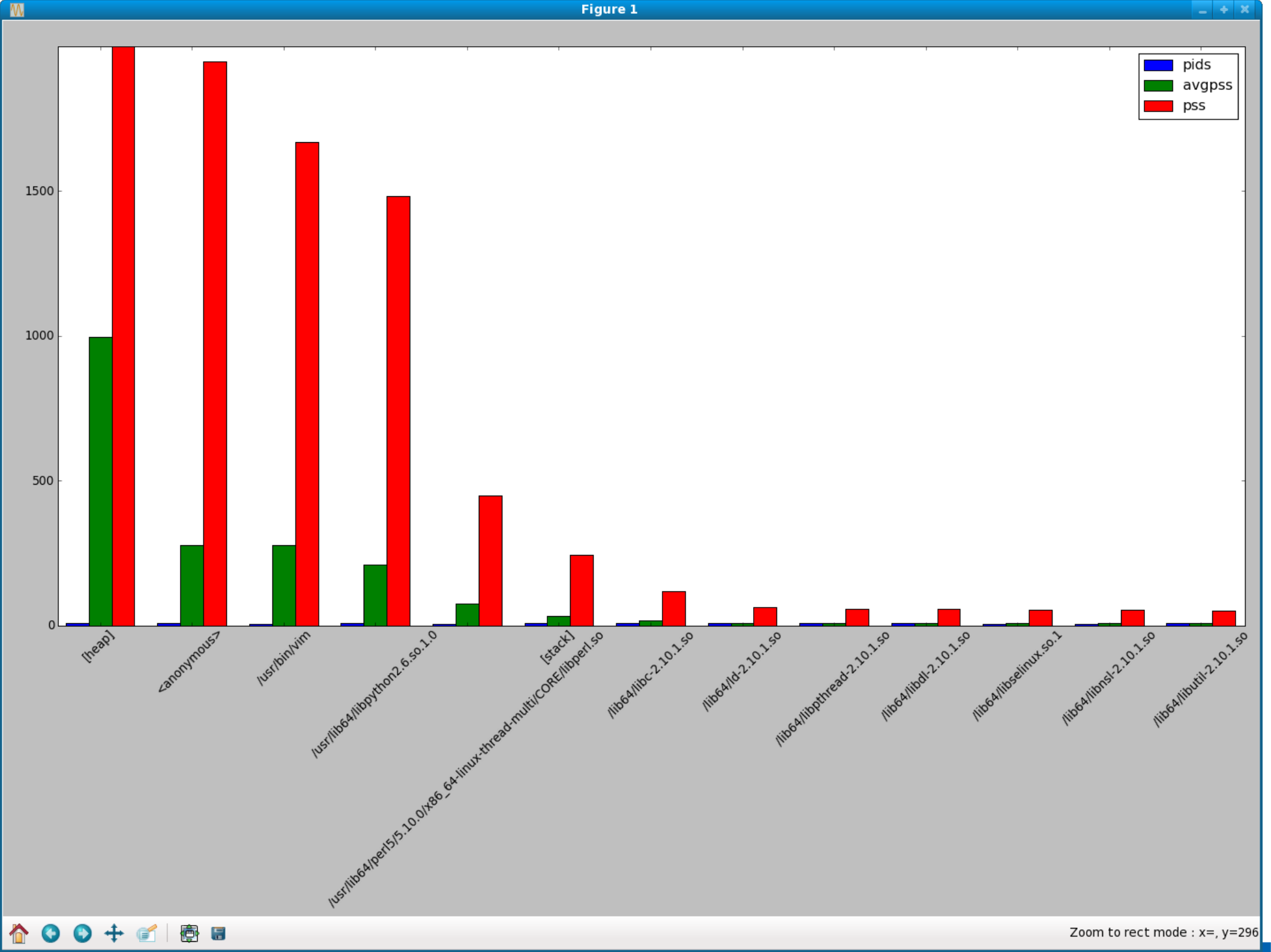
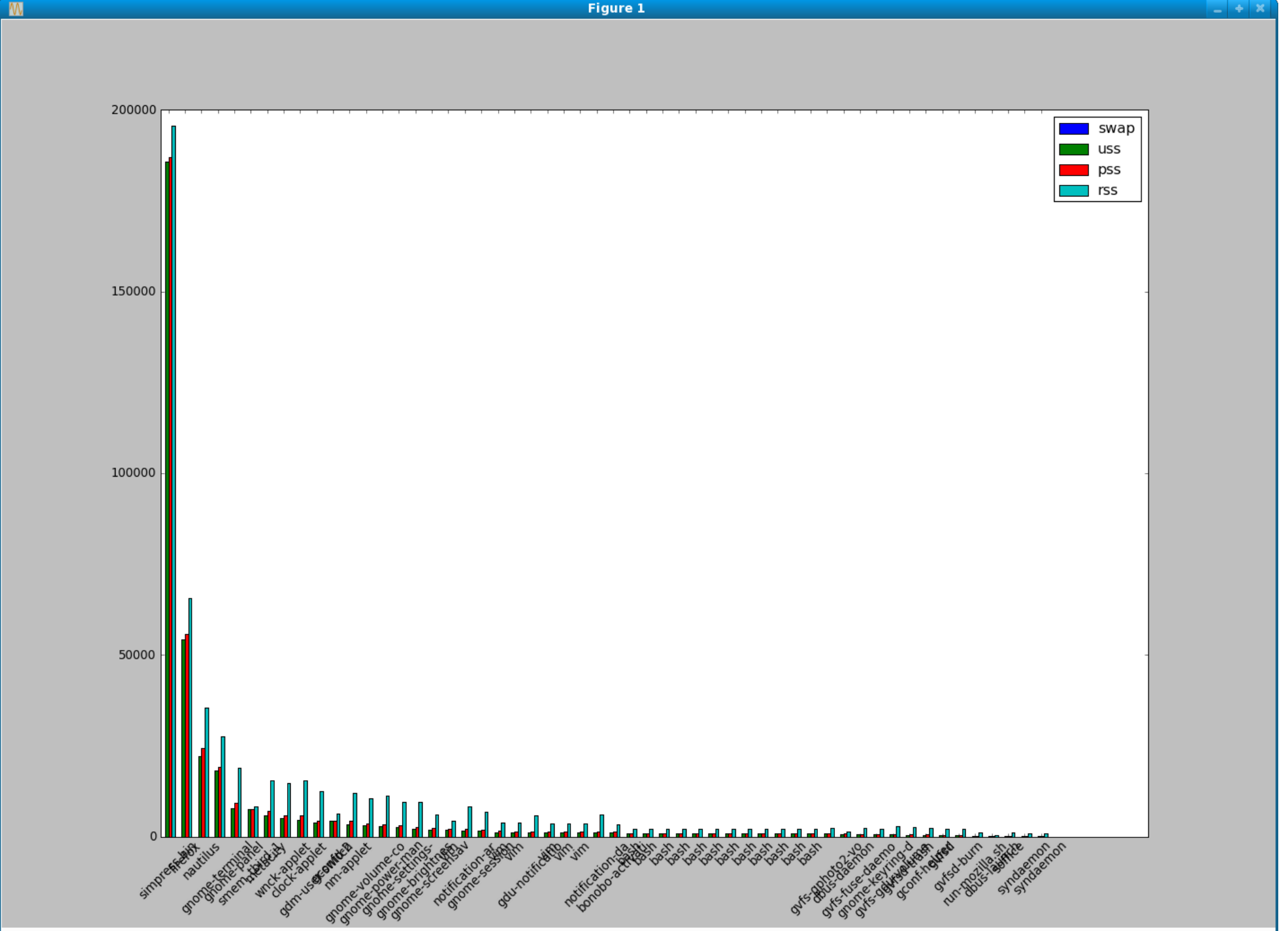


Figure 1
pids
avgpss
pss
1500
1000
500
0
[heap]
<anonymous>
/usr/bin/vim
/usr/lib64/libpython2.6.so.1.0
[stack]
/usr/lib64/perl5/5.10.0/x86_64-linux-thread-multi/CORE/libperl.so
/lib64/libc-2.10.1.so
/lib64/ld-2.10.1.so
/lib64/libpthread-2.10.1.so
/lib64/libdl-2.10.1.so
/lib64/libselinux.so.1
/lib64/libnsl-2.10.1.so
/lib64/libutil-2.10.1.so
Zoom to rect mode : x=, y=296

```
? smem -r
  PID User     Command                         Swap      USS      PSS      RSS
 6409 frowand  /usr/lib64/openoffice.org3/       0   185692   186954   195668
 6444 frowand  /usr/lib64/firefox-3.5b4/fi       0    54844    56432    66268
 2447 frowand  nautilus                          0    22028    24220    35168
 2530 frowand  gnome-terminal                    0    17412    18301    26752
 2428 frowand  gnome-panel                       0     7592     9005    18676
 6929 frowand  /usr/bin/python ./smem_tbir       0     7488     7510     8264
 2422 frowand  metacity                          0     5988     7011    15424
 2511 frowand  /usr/libexec/clock-applet -       0     4588     5866    15556
 2492 frowand  /usr/libexec/wnck-applet --       0     4812     5644    14436
 2515 frowand  /usr/libexec/gdm-user-switc       0     3988     4520    12632
 2413 frowand  /usr/libexec/gconfd-2             0     4376     4430     6308
 2462 frowand  nm-applet --sm-disable            0     3288     4352    12104
 2458 frowand  gnome-volume-control-applet       0     3248     3734    10648
 2455 frowand  gnome-power-manager               0     2876     3526    11216
 2417 frowand  /usr/libexec/gnome-settings       0     2772     3231     9552
 2513 frowand  /usr/libexec/gnome-brightne       0     2224     2731     9676
 2568 frowand  gnome-screensaver                 0     1944     2327     6172
 2602 frowand  vim outline                       0     1868     2177     4484
 2496 frowand  /usr/libexec/notification-a       0     1740     2081     8276
 2300 frowand  gnome-session                     0     1712     1932     6948
 6741 frowand  vim setup_fedora_11               0     1272     1578     3860
 2603 frowand  vim notes                         0     1236     1542     3844
 2451 frowand  /usr/libexec/gdu-notificati       0     1296     1517     5992

......
```

Figure 1


swap
uss
pss
rss
200000
150000
100000
50000
0
nautilus
wnck-applet
clock-applet
nm-applet
gvfsd-burn
syndaemon
Zoom to rect mode : x=, y=1.36e+05

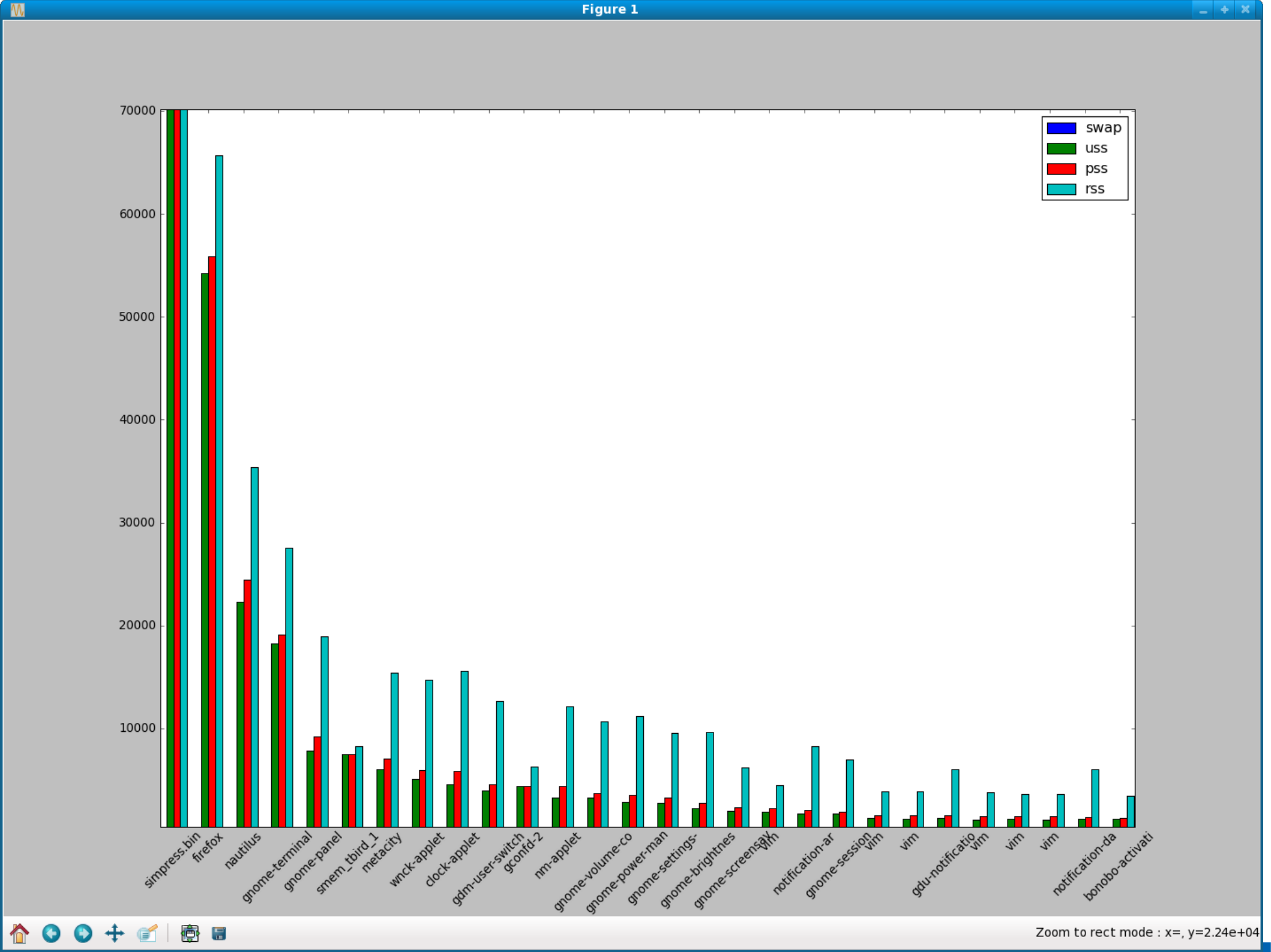


Figure 1
70000
60000
50000
40000
30000
20000
10000
swap
uss
pss
rss
simpress.bin
firefox
nautilus
gnome-terminal
gnome-panel
smem_tbird_1
metacity
wnck-applet
clock-applet
gdm-user-switch
gconfd-2
nm-applet
gnome-volume-co
gnome-power-man
gnome-settings-
gnome-brightnes
gnome-screensav
vim
notification-ar
gnome-session
vim
vim
gdu-notificatio
vim
vim
vim
notification-da
bonobo-activati
Zoom to rect mode : x=, y=2.24e+04

smem [memory]

Requires:

- kernel > 2.6.27
- Python 2.4
- python-matplotlib if charts are desired

oprofile [processor]

```
$ opcontrol --start --vmlinux=vmlinux-2.6.32-rc3.ftr
$ oprofile -l

samples  %        image name               app name                 symbol name
644900   22.2335  libc-2.10.1.so           libc-2.10.1.so           memcpy
481131   16.5874  no-vmlinux               no-vmlinux               /no-vmlinux
119529    4.1209  vmlinux                  vmlinux                  native_read_tsc
111184    3.8332  vmlinux                  vmlinux                  __lock_acquire
89747     3.0941  vmlinux                  vmlinux                  lock_release
69191     2.3854  vmlinux                  vmlinux                  lock_acquire
66124     2.2797  vmlinux                  vmlinux                  mwait_idle_with_hints
64803     2.2341  libpixman-1.so.0.14.0    libpixman-1.so.0.14.0    /usr/lib64/libpixman-1
59559     2.0534  vmlinux                  vmlinux                  read_hpet
53408     1.8413  vmlinux                  vmlinux                  native_sched_clock
45348     1.5634  vmlinux                  vmlinux                  lock_acquired
40988     1.4131  Xorg                     Xorg                     /usr/bin/Xorg
38490     1.3270  vmlinux                  vmlinux                  __d_lookup
33713     1.1623  libxul.so                libxul.so                /usr/lib64/xulrunner-1
33497     1.1548  vmlinux                  vmlinux                  lock_release_holdtime
28337     0.9769  vmlinux                  vmlinux                  __link_path_walk
27692     0.9547  ls                       ls                       /bin/ls
27472     0.9471  vmlinux                  vmlinux                  hlock_class
27261     0.9398  libglib-2.0.so.0.2000.1  libglib-2.0.so.0.2000.1  /lib64/libglib-2.0.so.
26236     0.9045  vmlinux                  vmlinux                  avc_has_perm_noaudit

......
```

oprofile [processor]

```
$ opcontrol --start --vmlinux=vmlinux-2.6.32-rc3.ftr
$ oprofile -l
```

```
samples %        image name      app name        symbol name
644900  22.2335 libc-2.10.1.so libc-2.10.1.so memcpy
481131  16.5874 no-vmlinux     no-vmlinux     /no-vmlinux
119529   4.1209 vmlinux        vmlinux        native_read_tsc
111184   3.8332 vmlinux        vmlinux        __lock_acquire
89747    3.0941 vmlinux        vmlinux        lock_release
69191    2.3854 vmlinux        vmlinux        lock_acquire
66124    2.2797 vmlinux        vmlinux        mwait_idle_with_h
64803    2.2341 libpixman-1.so libpixman-1.so /usr/lib64/libpix
```

oprofile [processor]

```
$ opcontrol --start --no-vmlinux
$ oprofile -l

samples  %        image name               app name                 symbol name
339274   69.6932  no-vmlinux               no-vmlinux               /no-vmlinux
18771     3.8559  Xorg                     Xorg                     /usr/bin/Xorg
13812     2.8372  libglib-2.0.so.0.2000.1  libglib-2.0.so.0.2000.1  /lib64/libglib-2.0.so.
13209     2.7134  libpixman-1.so.0.14.0    libpixman-1.so.0.14.0    /usr/lib64/libpixman-1
10744     2.2070  libcairo.so.2.10800.6    libcairo.so.2.10800.6    /usr/lib64/libcairo.so
7841      1.6107  radeon_drv.so            radeon_drv.so            /usr/lib64/xorg/module
6139      1.2611  libexa.so                libexa.so                /usr/lib64/xorg/module
6060      1.2448  libgobject-2.0.so.0.2000.1 libgobject-2.0.so.0.2000.1 /lib64/libgobject-
5178      1.0637  libvte.so.9.5.1          libvte.so.9.5.1          /usr/lib64/libvte.so.9
5083      1.0441  libpthread-2.10.1.so     libpthread-2.10.1.so     pthread_mutex_lock
3786      0.7777  libxul.so                libxul.so                /usr/lib64/xulrunner-1
3364      0.6910  libpthread-2.10.1.so     libpthread-2.10.1.so     pthread_mutex_unlock
3304      0.6787  libgdk-x11-2.0.so.0.1600.1 libgdk-x11-2.0.so.0.1600.1 /usr/lib64/libgdk-
3147      0.6465  libmozjs.so              libmozjs.so              /usr/lib64/xulrunner-1
2865      0.5885  libdbus-1.so.3.4.0       libdbus-1.so.3.4.0       /lib64/libdbus-1.so.3.
2703      0.5552  libc-2.10.1.so           libc-2.10.1.so           _int_malloc
2190      0.4499  libX11.so.6.2.0          libX11.so.6.2.0          /usr/lib64/libX11.so.6
1945      0.3995  libc-2.10.1.so           libc-2.10.1.so           memcpy
1868      0.3837  libc-2.10.1.so           libc-2.10.1.so           _int_free
1861      0.3823  libgtk-x11-2.0.so.0.1600.1 libgtk-x11-2.0.so.0.1600.1 /usr/lib64/libgtk-

......
```

oprofile [processor]

- Time Based Sampling (possible side effects)
- How often is a function executing?

systemtap [processor]

Example from Fedora fc11

/usr/share/doc/systemtap-0.9.5/examples/process/syscalltimes.txt

```
$ ./syscalltimes -n top -n vi
   System Call Count Total ns Avg ns Min ns Max ns
        access     4    59169  14792   7770  27556
rt_sigprocmask     6    32879   5479   5328   6132
  rt_sigaction    22   119127   5414   5291   6293
       sysinfo     1    13178  13178  13178  13178
         lseek     2    12869   6434   6331   6538
        getuid     3    16839   5613   5468   5757
     getrlimit     1     5558   5558   5558   5558
        munmap     8   103326  12915   9689  19190
        getpid     1     5639   5639   5639   5639
        unlink     3   143502  47834  20291  97191
```

systemtap [processor]

- Debugging tool, but previous example shows example of calculating system call duration

time [processor]

Do not always need the new, powerful tools.

Example usage:

Compare two algorithms

- elapsed time
- user space time
- kernel space time

time [processor]

```
$ time ./test_array_1_64 >/dev/null

real  0m1.968s
user  0m1.946s
sys   0m0.022s

$ time ./test_array_2_64 >/dev/null

real  0m8.069s
user  0m8.038s
sys   0m0.022s
```

Performance Events [processor]

Do not always need the new, powerful tools.

But sometimes the new tools give new capability.

Performance Events [processor]

```
$ perf stat --repeat 10 ./test_array_2_64 >/dev/null

 Performance counter stats for './test_array_2_64' (10 runs):

    7898.602695  task-clock-msecs         #      0.994 CPUs    ( +-   0.391% )
            810  context-switches         #      0.000 M/sec   ( +-   1.024% )
              0  CPU-migrations           #      0.000 M/sec   ( +-  66.667% )
           2156  page-faults              #      0.000 M/sec   ( +-   0.000% )
    19811028275  cycles                   #   2508.169 M/sec   ( +-   0.402% )
     5913651907  instructions             #      0.299 IPC     ( +-   0.006% )
      532668283  cache-references         #     67.438 M/sec   ( +-   0.017% )
      174705853  cache-misses             #     22.119 M/sec   ( +-   0.944% )

    7.946727335  seconds time elapsed   ( +-   0.389% )
```

Performance Events [processor]

test_array_1_64 vs. test_array_2_64

```
         241  context-switches #        0.000 M/sec
         810  context-switches #        0.000 M/sec

  4874649281  cycles           #     2506.837 M/sec
 19811028275  cycles           #     2508.169 M/sec

  5836419424  instructions     #        1.197 IPC
  5913651907  instructions     #        0.299 IPC

    66860331  cache-references #       34.384 M/sec
   532668283  cache-references #       67.438 M/sec

     1219420  cache-misses     #        0.627 M/sec
   174705853  cache-misses     #       22.119 M/sec

1.958908427  seconds time elapsed
7.946727335  seconds time elapsed
```

Performance Events [processor]

```
/* test_array_1.c */
#define A_SIZE 1024
long array [A_SIZE][A_SIZE];

total = 0;

    for (j=0; j < A_SIZE; j++)
        for (k=0; k < A_SIZE; k++) {
            array[j][k] = j + k;
            total += array[j][k];
        }

for (i=0; i < 500; i++)
    for (j=0; j < A_SIZE; j++)
        for (k=0; k < A_SIZE; k++)
            total += array[j][k];
```

Performance Events [processor]

```
diff test_array_1.c test_array_2.c

-               array[j][k] = j + k;
+               array[k][j] = j + k;

-               total += array[j][k];
+               total += array[k][j];

-               total += array[j][k];
+               total += array[k][j];
```

bootchart [timeline]

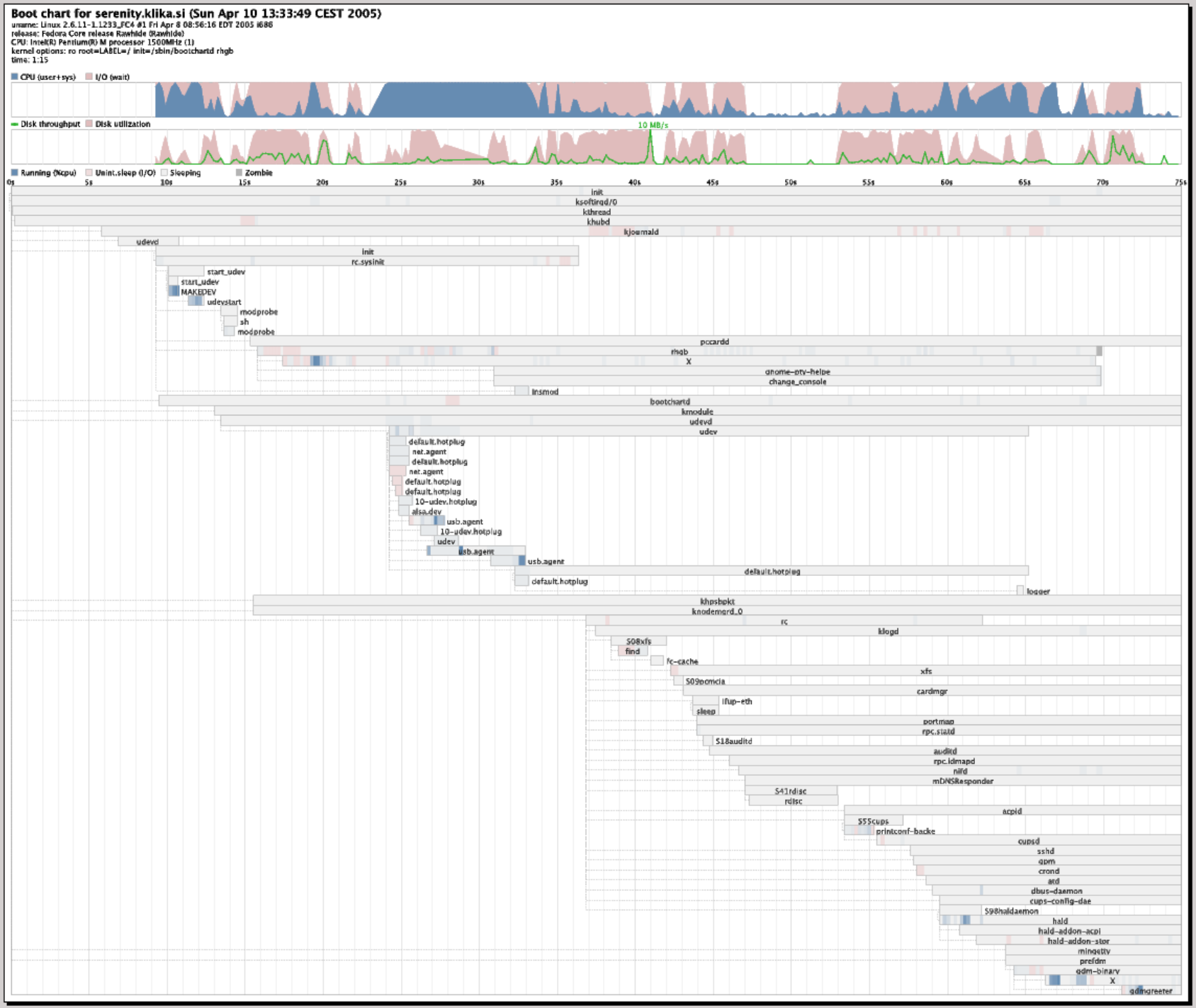
- After kernel boot, the init phase
- The following images are from

  http://www.bootchart.org/

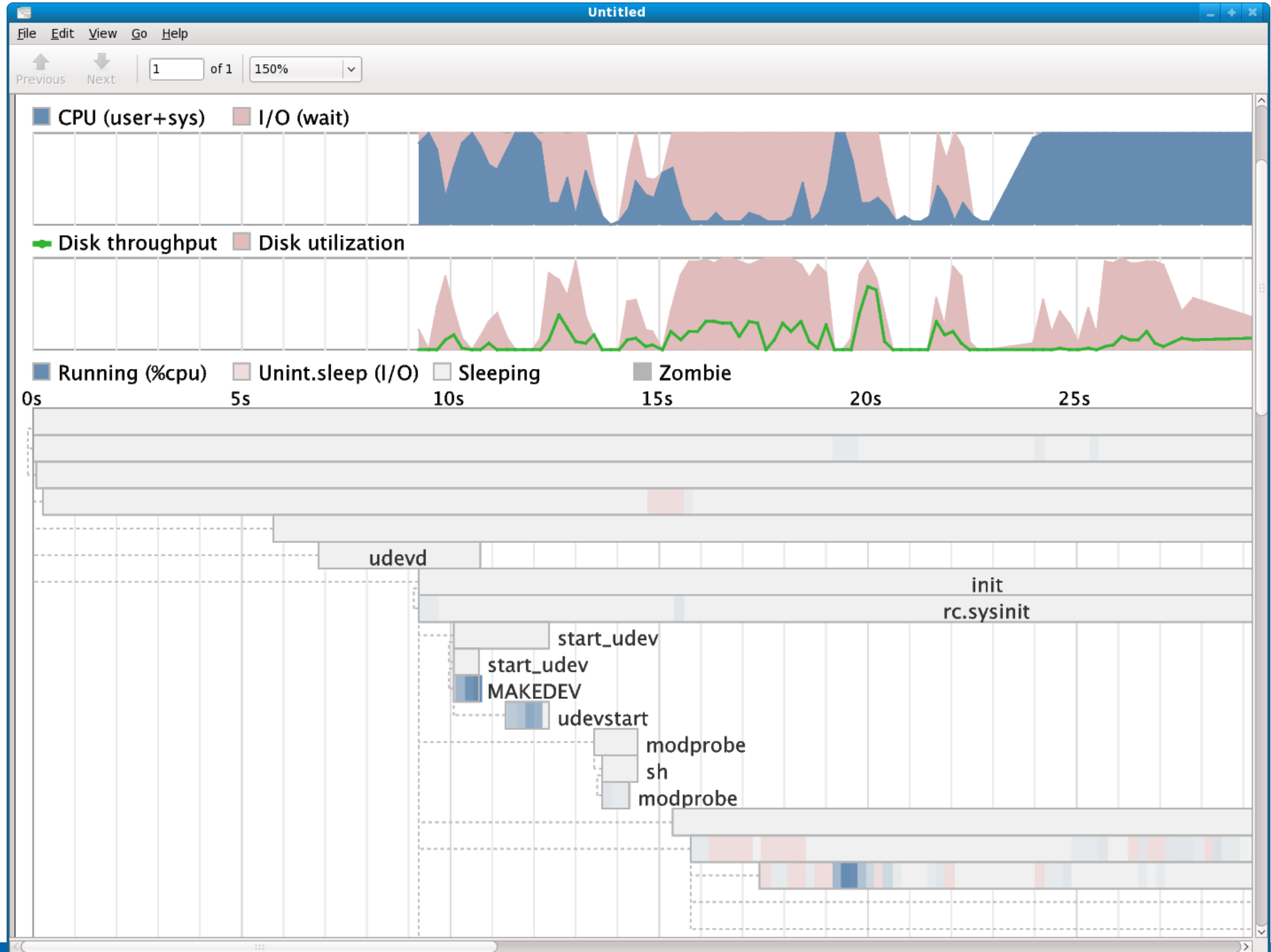
Untitled

File Edit View Go Help

Previous Next 1 of 1 Best Fit



Untitled
File Edit View Go Help
Previous Next
1 of 1
125%
Boot chart for serenity.klika.si (Sun Apr 10 13:33:49 CEST 2005)
uname: Linux 2.6.11-1.1233_FC4 #1 Fri Apr 8 08:56:16 EDT 2005 i686
release: Fedora Core release Rawhide (Rawhide)
CPU: Intel(R) Pentium(R) M processor 1500MHz (1)
kernel options: ro root=LABEL=/ init=/sbin/bootchartd rhgb
time: 1:15
CPU (user+sys)
I/O (wait)
Disk throughput
Disk utilization
Running (%cpu)
Unint.sleep (I/O)
Sleeping
Zombie
0s
5s
10s
15s
20s
25s
30s
udevd
init
rc.sysinit
start_udev
start_udev
MAKEDEV
udevstart
modprobe
sh
modprobe


Untitled
File Edit View Go Help
Previous Next
1 of 1
150%
CPU (user+sys)
I/O (wait)
Disk throughput
Disk utilization
Running (%cpu)
Unint.sleep (I/O)
Sleeping
Zombie
0s
5s
10s
15s
20s
25s
udevd
init
rc.sysinit
start_udev
start_udev
MAKEDEV
udevstart
modprobe
sh
modprobe

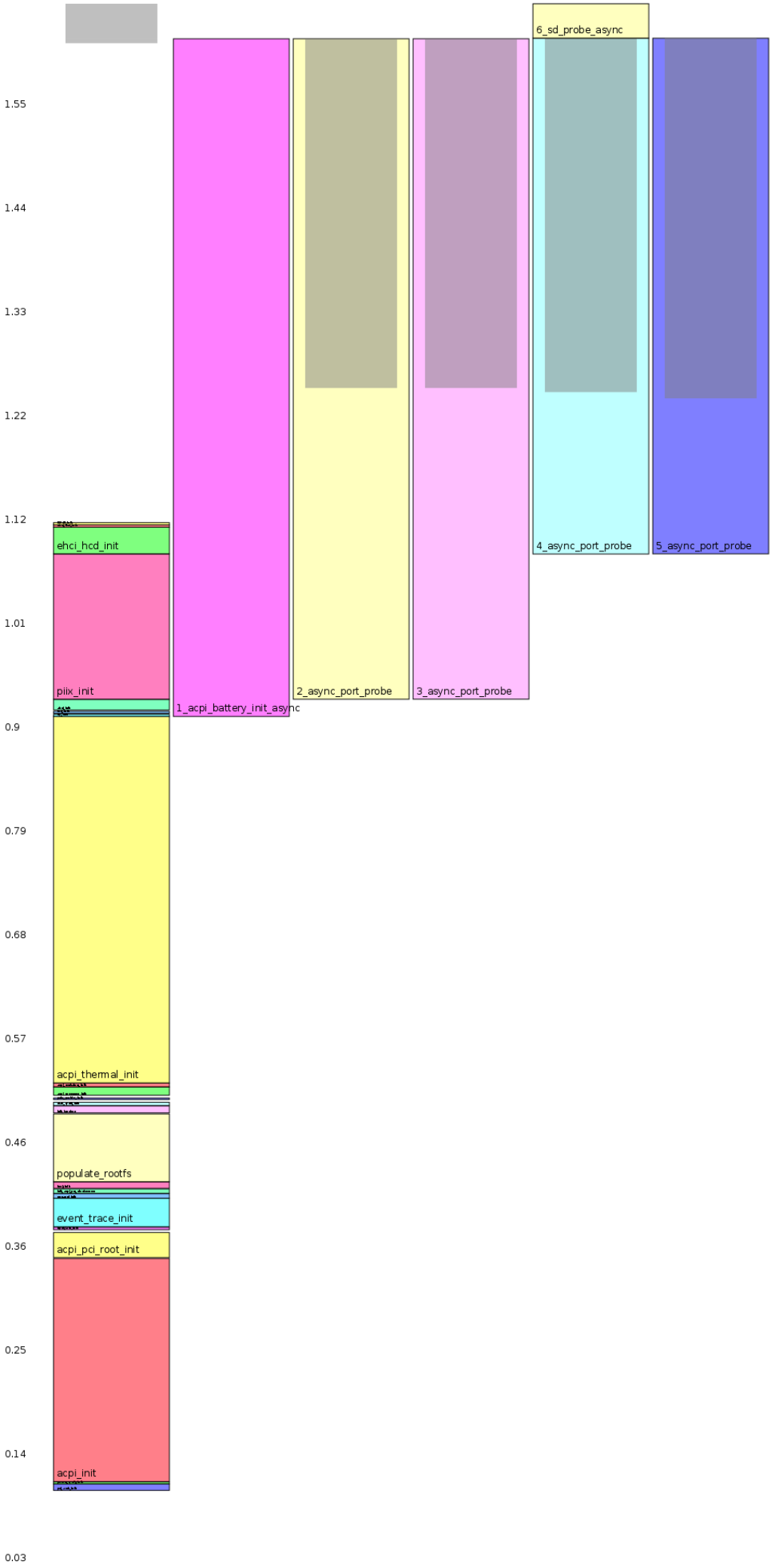
boot tracer [timeline]

- Late kernel boot: initcall functions
- Data source: dmesg, printk times:

```
[0.107274] calling  net_ns_init
[0.107323] initcall net_ns_init
[0.107327] calling  e820_mark_nvs_memory
[0.107335] initcall e820_mark_nvs_memory
[0.107338] calling  cpufreq_tsc
[0.107341] initcall cpufreq_tsc
[0.107345] calling  pci_reboot_init
[0.107348] initcall pci_reboot_init
```

boot tracer [timeline]

```
# enable CONFIG_BOOT_TRACER
# and build kernel

# boot with the parameters:
   "initcall_debug" and "printk.time=1"

# scripts is in the kernel source tree
dmesg | perl scripts/bootgraph.pl > output.svg

inkscape output.svg
```


6_sd_probe_async
1.55
1.44
1.33
1.22
1.12
1.01
0.9
0.79
0.68
0.57
0.46
0.36
0.25
0.14
0.03
ehci_hcd_init
piix_init
1_acpi_battery_init_async
2_async_port_probe
3_async_port_probe
4_async_port_probe
5_async_port_probe
acpi_thermal_init
populate_rootfs
event_trace_init
acpi_pci_root_init
acpi_init

boot tracer [timeline]

```
Requires:

    -kernel >= 2.6.28
```

kft [timeline]
Kernel Function Trace

- Early kernel boot: start_kernel()
- Can also be used after boot.
- Triggers and filters to reduce amount of data logged.
- Post trace analysis tool does filtering, sorting, analysis and trace formatting.
- Can provide function level detail.

kft [timeline]
Kernel Function Trace

- Previously known as KFI, Kernel Function Instrumentation

- Patches available for 2.6.7 - 2.6.28

kft [timeline]
Kernel Function Trace

Tim Bird is moving this technology forward into the ftrace framework. See his presentation from yesterday:

Analyzing Kernel Function Execution with ftrace

lttng

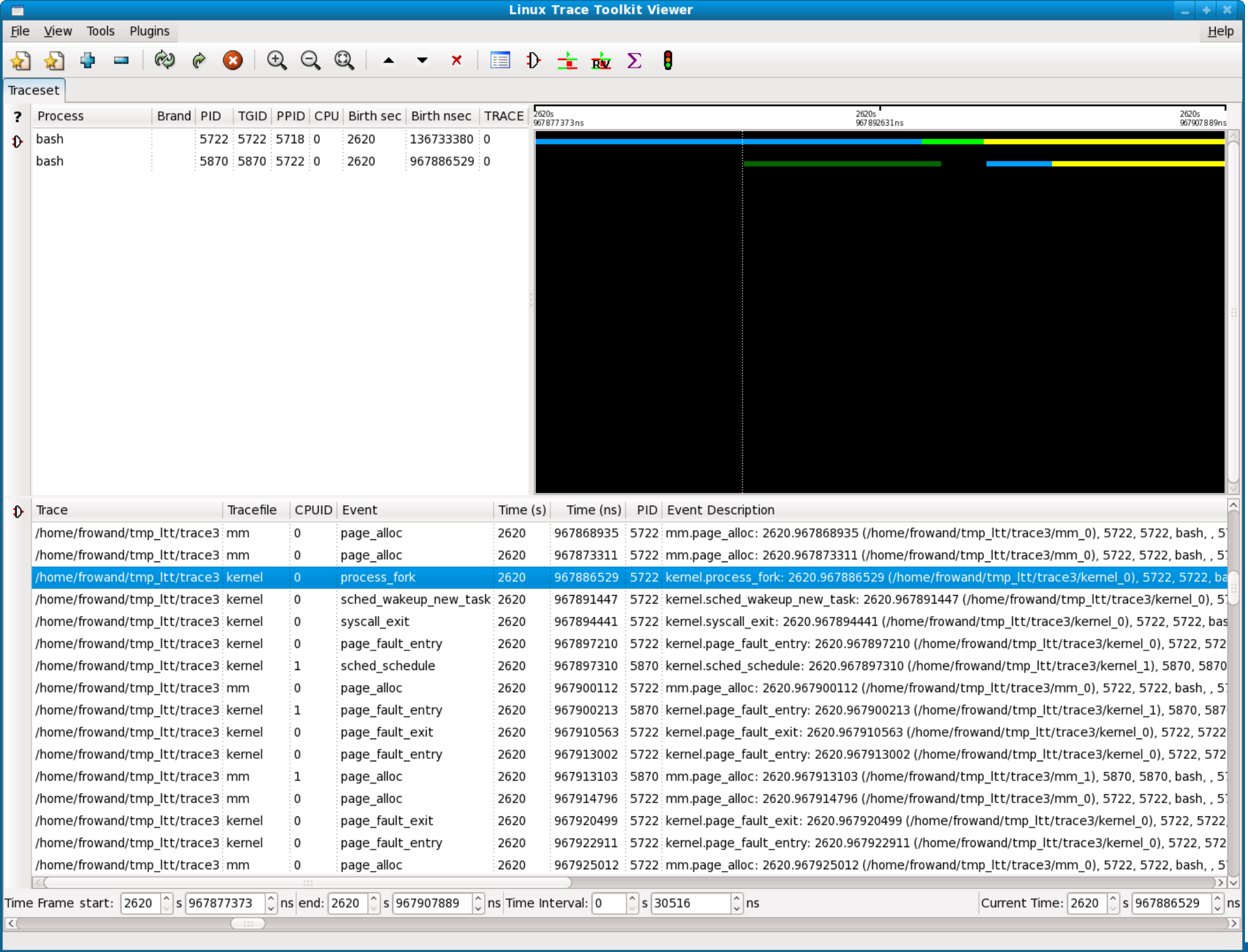
[timeline]

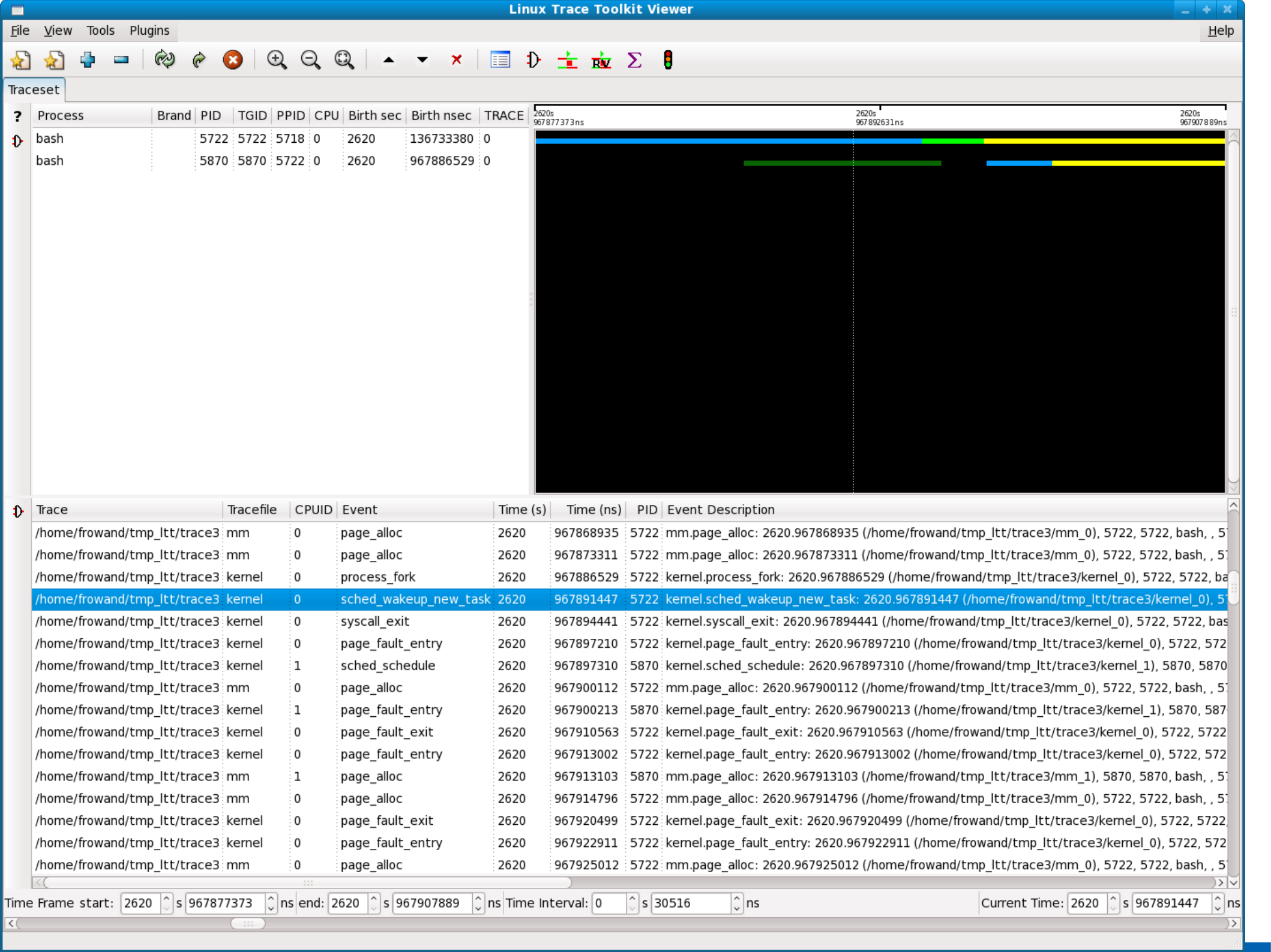
Long history, starting with predecessor ltt.

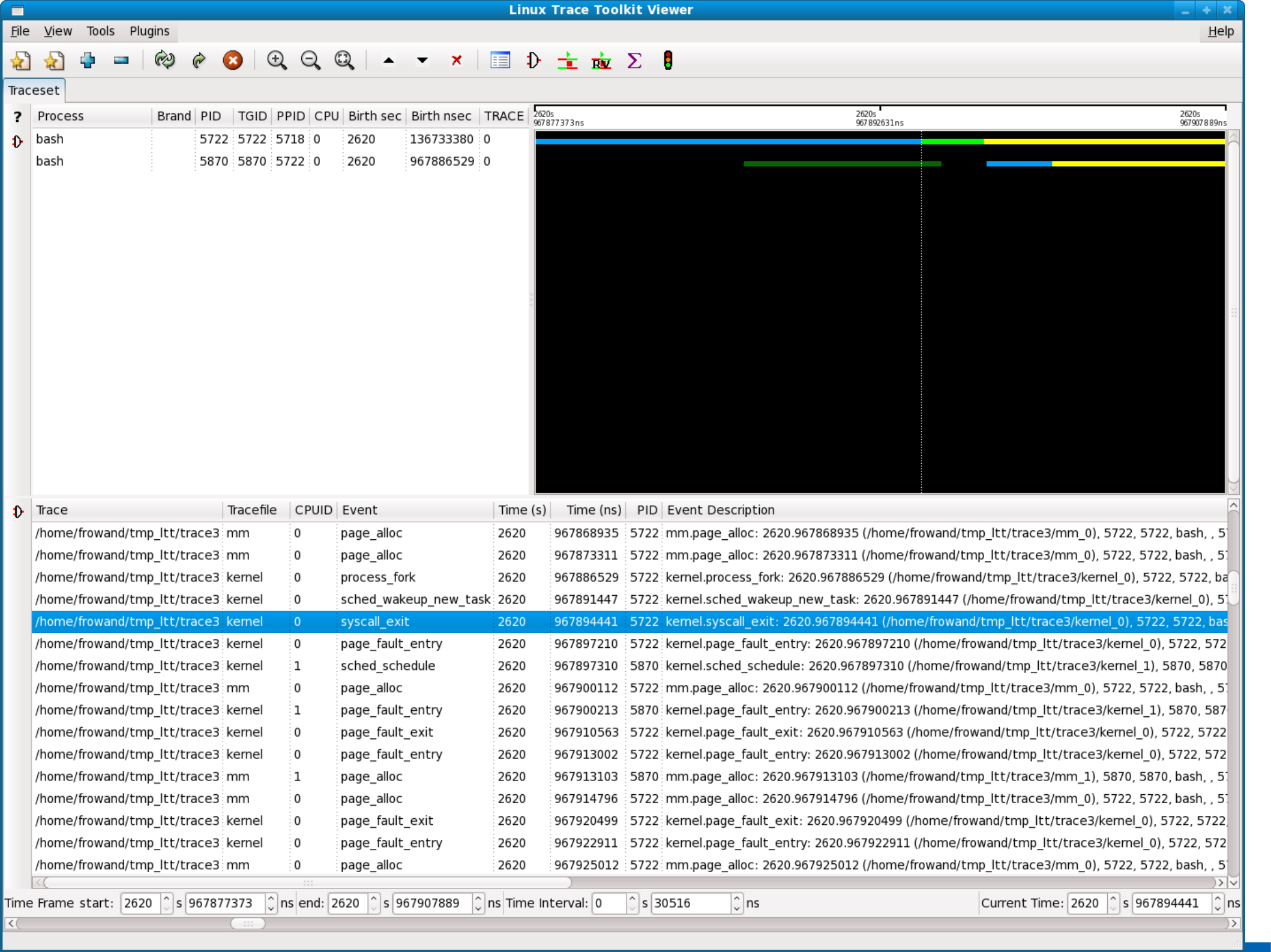
Not in mainline.

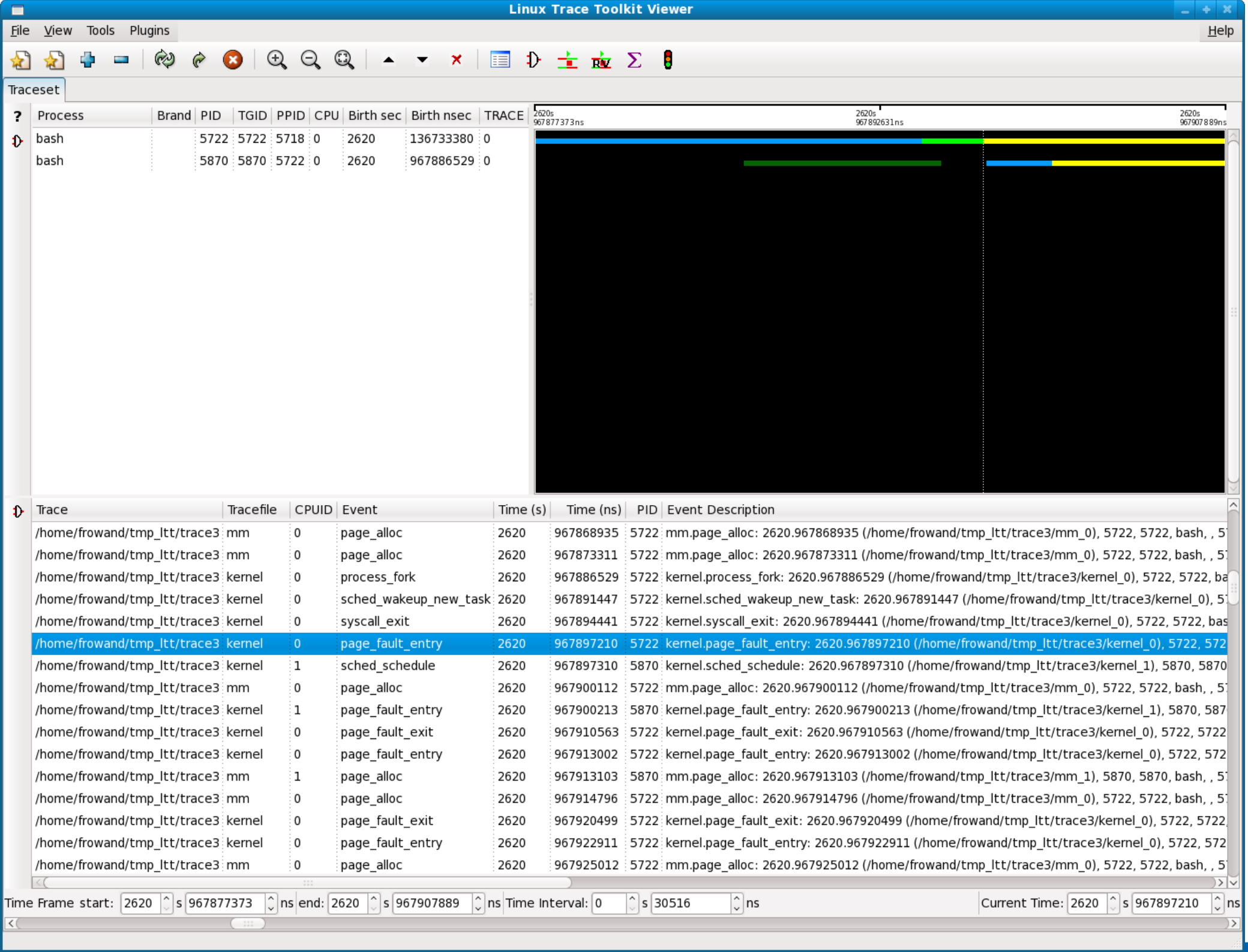
Included in some distributions:

- MontaVista
- Wind River


Linux Trace Toolkit Viewer
File View Tools Plugins Help
Traceset
Process Brand PID TGID PPID CPU Birth sec Birth nsec TRACE
bash 5722 5722 5718 0 2620 136733380 0
bash 5870 5870 5722 0 2620 967886529 0
2620s 967877373ns
2620s 967892631ns
2620s 967907889ns
Trace Tracefile CPUID Event Time (s) Time (ns) PID Event Description
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967868935 5722 mm.page_alloc: 2620.967868935 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967873311 5722 mm.page_alloc: 2620.967873311 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 0 process_fork 2620 967886529 5722 kernel.process_fork: 2620.967886529 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, ba
/home/frowand/tmp_ltt/trace3 kernel 0 sched_wakeup_new_task 2620 967891447 5722 kernel.sched_wakeup_new_task: 2620.967891447 (/home/frowand/tmp_ltt/trace3/kernel_0), 5
/home/frowand/tmp_ltt/trace3 kernel 0 syscall_exit 2620 967894441 5722 kernel.syscall_exit: 2620.967894441 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, bas
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967897210 5722 kernel.page_fault_entry: 2620.967897210 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 kernel 1 sched_schedule 2620 967897310 5870 kernel.sched_schedule: 2620.967897310 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 5870
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967900112 5722 mm.page_alloc: 2620.967900112 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 1 page_fault_entry 2620 967900213 5870 kernel.page_fault_entry: 2620.967900213 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 587
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_exit 2620 967910563 5722 kernel.page_fault_exit: 2620.967910563 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967913002 5722 kernel.page_fault_entry: 2620.967913002 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 mm 1 page_alloc 2620 967913103 5870 mm.page_alloc: 2620.967913103 (/home/frowand/tmp_ltt/trace3/mm_1), 5870, 5870, bash, , 5
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967914796 5722 mm.page_alloc: 2620.967914796 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_exit 2620 967920499 5722 kernel.page_fault_exit: 2620.967920499 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722,
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967922911 5722 kernel.page_fault_entry: 2620.967922911 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967925012 5722 mm.page_alloc: 2620.967925012 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
Time Frame start: 2620 s 967877373 ns end: 2620 s 967907889 ns Time Interval: 0 s 30516 ns
Current Time: 2620 s 967886529 ns


Linux Trace Toolkit Viewer
File View Tools Plugins
Help
Traceset
Process Brand PID TGID PPID CPU Birth sec Birth nsec TRACE
bash 5722 5722 5718 0 2620 136733380 0
bash 5870 5870 5722 0 2620 967886529 0
2620s 967877373ns
2620s 967892631ns
2620s 967907889ns
Trace Tracefile CPUID Event Time (s) Time (ns) PID Event Description
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967868935 5722 mm.page_alloc: 2620.967868935 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967873311 5722 mm.page_alloc: 2620.967873311 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 0 process_fork 2620 967886529 5722 kernel.process_fork: 2620.967886529 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, ba
/home/frowand/tmp_ltt/trace3 kernel 0 sched_wakeup_new_task 2620 967891447 5722 kernel.sched_wakeup_new_task: 2620.967891447 (/home/frowand/tmp_ltt/trace3/kernel_0), 5
/home/frowand/tmp_ltt/trace3 kernel 0 syscall_exit 2620 967894441 5722 kernel.syscall_exit: 2620.967894441 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, bas
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967897210 5722 kernel.page_fault_entry: 2620.967897210 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 kernel 1 sched_schedule 2620 967897310 5870 kernel.sched_schedule: 2620.967897310 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 5870
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967900112 5722 mm.page_alloc: 2620.967900112 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 1 page_fault_entry 2620 967900213 5870 kernel.page_fault_entry: 2620.967900213 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 587
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_exit 2620 967910563 5722 kernel.page_fault_exit: 2620.967910563 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967913002 5722 kernel.page_fault_entry: 2620.967913002 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 mm 1 page_alloc 2620 967913103 5870 mm.page_alloc: 2620.967913103 (/home/frowand/tmp_ltt/trace3/mm_1), 5870, 5870, bash, , 5
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967914796 5722 mm.page_alloc: 2620.967914796 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_exit 2620 967920499 5722 kernel.page_fault_exit: 2620.967920499 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722,
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967922911 5722 kernel.page_fault_entry: 2620.967922911 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967925012 5722 mm.page_alloc: 2620.967925012 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
Time Frame start: 2620 s 967877373 ns end: 2620 s 967907889 ns Time Interval: 0 s 30516 ns
Current Time: 2620 s 967891447 ns


Linux Trace Toolkit Viewer
File View Tools Plugins Help
Traceset
Process Brand PID TGID PPID CPU Birth sec Birth nsec TRACE
bash 5722 5722 5718 0 2620 136733380 0
bash 5870 5870 5722 0 2620 967886529 0
2620s 967877373ns
2620s 967892631ns
2620s 967907889ns
Trace Tracefile CPUID Event Time (s) Time (ns) PID Event Description
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967868935 5722 mm.page_alloc: 2620.967868935 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967873311 5722 mm.page_alloc: 2620.967873311 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 0 process_fork 2620 967886529 5722 kernel.process_fork: 2620.967886529 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, ba
/home/frowand/tmp_ltt/trace3 kernel 0 sched_wakeup_new_task 2620 967891447 5722 kernel.sched_wakeup_new_task: 2620.967891447 (/home/frowand/tmp_ltt/trace3/kernel_0), 5
/home/frowand/tmp_ltt/trace3 kernel 0 syscall_exit 2620 967894441 5722 kernel.syscall_exit: 2620.967894441 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, bas
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967897210 5722 kernel.page_fault_entry: 2620.967897210 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 kernel 1 sched_schedule 2620 967897310 5870 kernel.sched_schedule: 2620.967897310 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 5870
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967900112 5722 mm.page_alloc: 2620.967900112 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 1 page_fault_entry 2620 967900213 5870 kernel.page_fault_entry: 2620.967900213 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 587
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_exit 2620 967910563 5722 kernel.page_fault_exit: 2620.967910563 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967913002 5722 kernel.page_fault_entry: 2620.967913002 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 mm 1 page_alloc 2620 967913103 5870 mm.page_alloc: 2620.967913103 (/home/frowand/tmp_ltt/trace3/mm_1), 5870, 5870, bash, , 5
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967914796 5722 mm.page_alloc: 2620.967914796 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_exit 2620 967920499 5722 kernel.page_fault_exit: 2620.967920499 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722,
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967922911 5722 kernel.page_fault_entry: 2620.967922911 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967925012 5722 mm.page_alloc: 2620.967925012 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
Time Frame start: 2620 s 967877373 ns end: 2620 s 967907889 ns Time Interval: 0 s 30516 ns
Current Time: 2620 s 967894441 ns

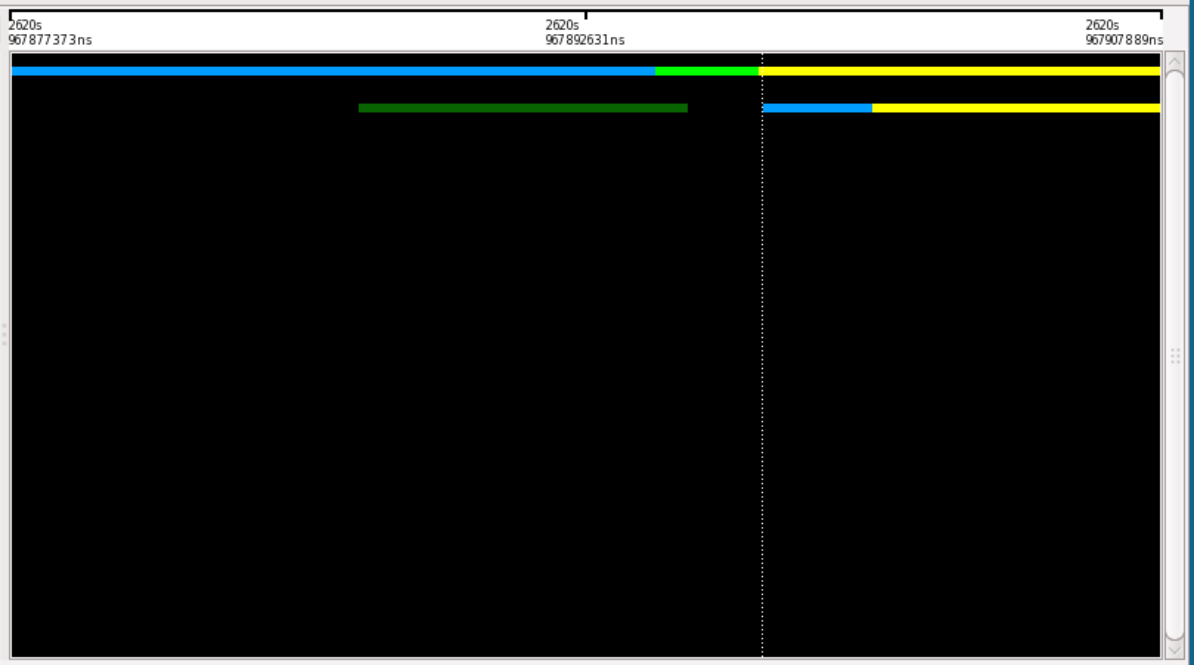

Linux Trace Toolkit Viewer
File View Tools Plugins Help
Traceset
Process Brand PID TGID PPID CPU Birth sec Birth nsec TRACE
bash 5722 5722 5718 0 2620 136733380 0
bash 5870 5870 5722 0 2620 967886529 0
2620s 967877373ns
2620s 967892631ns
2620s 967907889ns
Trace Tracefile CPUID Event Time (s) Time (ns) PID Event Description
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967868935 5722 mm.page_alloc: 2620.967868935 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967873311 5722 mm.page_alloc: 2620.967873311 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 0 process_fork 2620 967886529 5722 kernel.process_fork: 2620.967886529 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, ba
/home/frowand/tmp_ltt/trace3 kernel 0 sched_wakeup_new_task 2620 967891447 5722 kernel.sched_wakeup_new_task: 2620.967891447 (/home/frowand/tmp_ltt/trace3/kernel_0), 5
/home/frowand/tmp_ltt/trace3 kernel 0 syscall_exit 2620 967894441 5722 kernel.syscall_exit: 2620.967894441 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, bas
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967897210 5722 kernel.page_fault_entry: 2620.967897210 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 kernel 1 sched_schedule 2620 967897310 5870 kernel.sched_schedule: 2620.967897310 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 5870
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967900112 5722 mm.page_alloc: 2620.967900112 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 1 page_fault_entry 2620 967900213 5870 kernel.page_fault_entry: 2620.967900213 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 587
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_exit 2620 967910563 5722 kernel.page_fault_exit: 2620.967910563 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967913002 5722 kernel.page_fault_entry: 2620.967913002 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 mm 1 page_alloc 2620 967913103 5870 mm.page_alloc: 2620.967913103 (/home/frowand/tmp_ltt/trace3/mm_1), 5870, 5870, bash, , 5
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967914796 5722 mm.page_alloc: 2620.967914796 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_exit 2620 967920499 5722 kernel.page_fault_exit: 2620.967920499 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722,
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967922911 5722 kernel.page_fault_entry: 2620.967922911 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967925012 5722 mm.page_alloc: 2620.967925012 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
Time Frame start: 2620 s 967877373 ns end: 2620 s 967907889 ns Time Interval: 0 s 30516 ns
Current Time: 2620 s 967897210 ns

Linux Trace Toolkit Viewer

File View Tools Plugins Help

Traceset

Process	Brand	PID	TGID	PPID	CPU	Birth sec	Birth nsec	TRACE
bash		5722	5722	5718	0	2620	136733380	0
bash		5870	5870	5722	0	2620	967886529	0



Trace	Tracefile	CPUID	Event	Time (s)	Time (ns)	PID	Event Description
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967868935	5722	mm.page_alloc: 2620.967868935 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967873311	5722	mm.page_alloc: 2620.967873311 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3	kernel	0	process_fork	2620	967886529	5722	kernel.process_fork: 2620.967886529 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, ba
/home/frowand/tmp_ltt/trace3	kernel	0	sched_wakeup_new_task	2620	967891447	5722	kernel.sched_wakeup_new_task: 2620.967891447 (/home/frowand/tmp_ltt/trace3/kernel_0), 5
/home/frowand/tmp_ltt/trace3	kernel	0	syscall_exit	2620	967894441	5722	kernel.syscall_exit: 2620.967894441 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, bas
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_entry	2620	967897210	5722	kernel.page_fault_entry: 2620.967897210 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3	kernel	1	sched_schedule	2620	967897310	5870	kernel.sched_schedule: 2620.967897310 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 5870
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967900112	5722	mm.page_alloc: 2620.967900112 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3	kernel	1	page_fault_entry	2620	967900213	5870	kernel.page_fault_entry: 2620.967900213 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 587
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_exit	2620	967910563	5722	kernel.page_fault_exit: 2620.967910563 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_entry	2620	967913002	5722	kernel.page_fault_entry: 2620.967913002 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3	mm	1	page_alloc	2620	967913103	5870	mm.page_alloc: 2620.967913103 (/home/frowand/tmp_ltt/trace3/mm_1), 5870, 5870, bash, , 5
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967914796	5722	mm.page_alloc: 2620.967914796 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_exit	2620	967920499	5722	kernel.page_fault_exit: 2620.967920499 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722,
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_entry	2620	967922911	5722	kernel.page_fault_entry: 2620.967922911 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967925012	5722	mm.page_alloc: 2620.967925012 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5

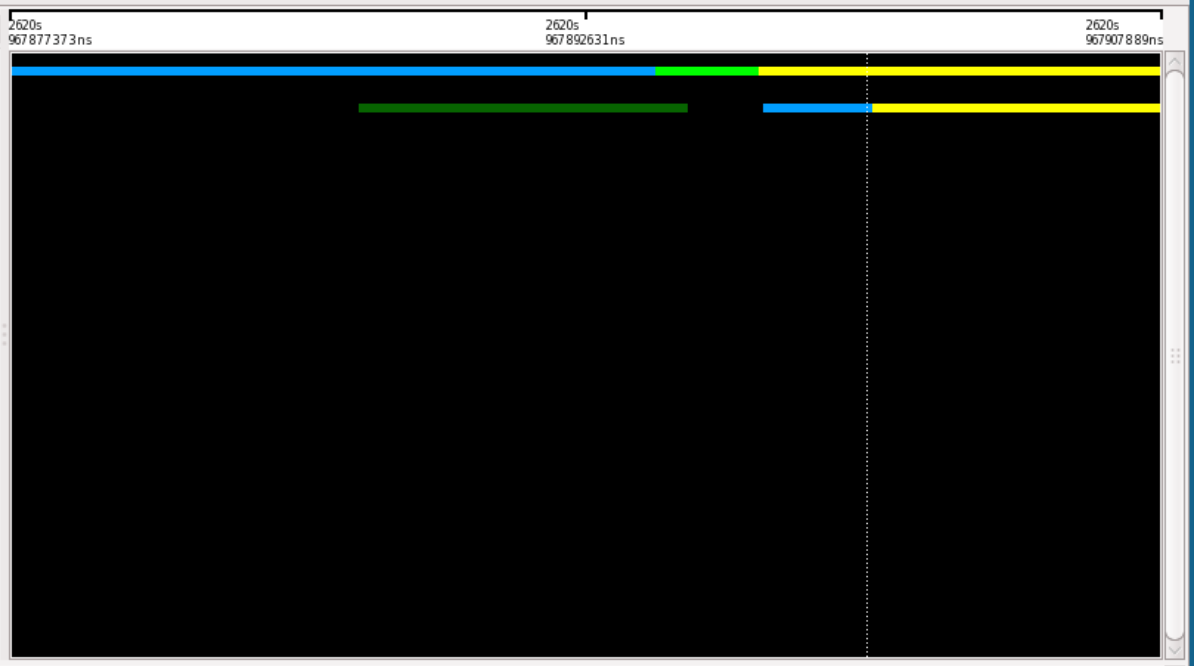
Time Frame start: 2620 s 967877373 ns end: 2620 s 967907889 ns Time Interval: 0 s 30516 ns Current Time: 2620 s 967897310 ns

Linux Trace Toolkit Viewer

File View Tools Plugins Help

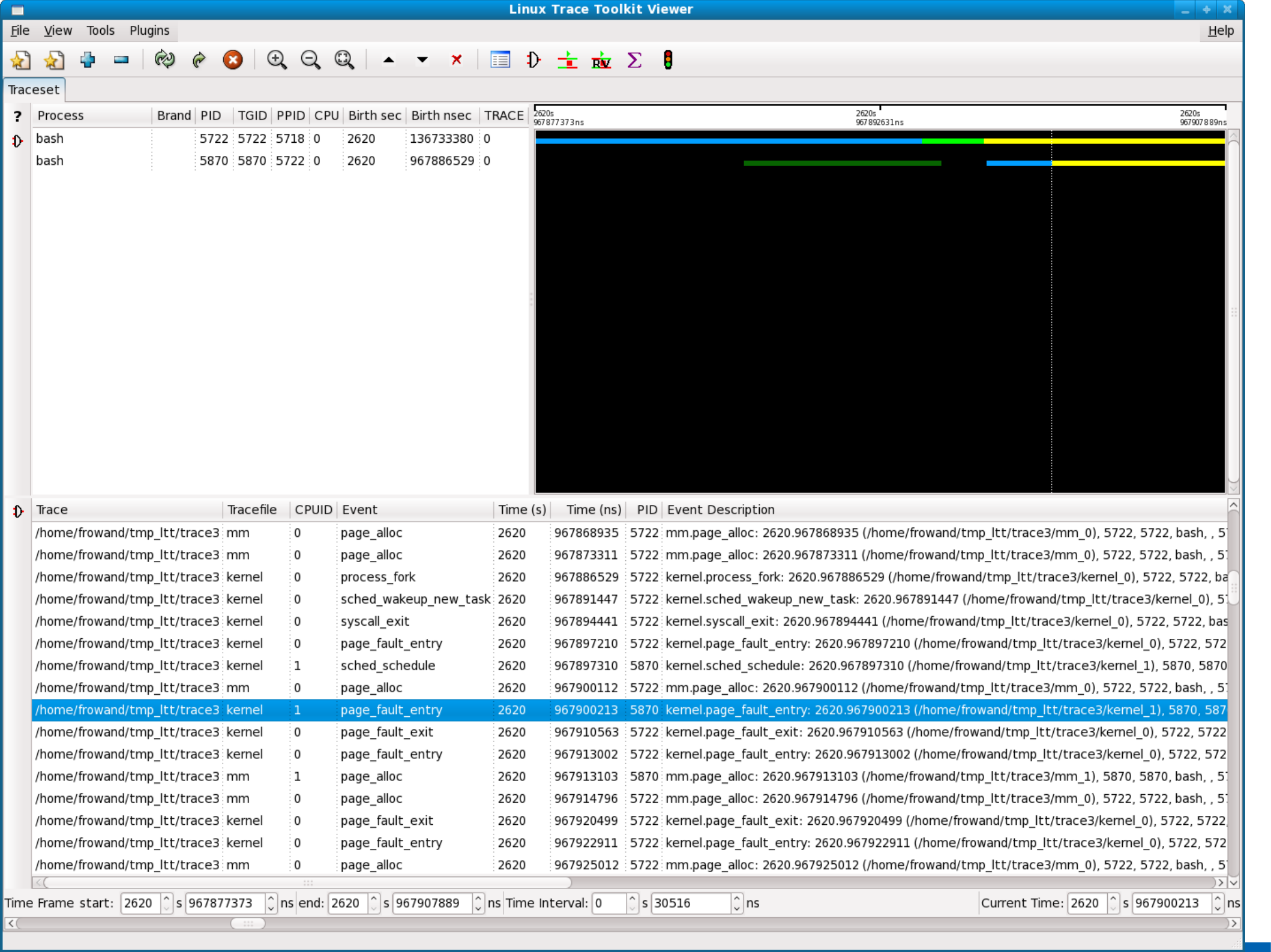
Traceset

Process	Brand	PID	TGID	PPID	CPU	Birth sec	Birth nsec	TRACE
bash		5722	5722	5718	0	2620	136733380	0
bash		5870	5870	5722	0	2620	967886529	0



Trace	Tracefile	CPUID	Event	Time (s)	Time (ns)	PID	Event Description
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967868935	5722	mm.page_alloc: 2620.967868935 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967873311	5722	mm.page_alloc: 2620.967873311 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3	kernel	0	process_fork	2620	967886529	5722	kernel.process_fork: 2620.967886529 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, ba
/home/frowand/tmp_ltt/trace3	kernel	0	sched_wakeup_new_task	2620	967891447	5722	kernel.sched_wakeup_new_task: 2620.967891447 (/home/frowand/tmp_ltt/trace3/kernel_0), 5
/home/frowand/tmp_ltt/trace3	kernel	0	syscall_exit	2620	967894441	5722	kernel.syscall_exit: 2620.967894441 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, bas
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_entry	2620	967897210	5722	kernel.page_fault_entry: 2620.967897210 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3	kernel	1	sched_schedule	2620	967897310	5870	kernel.sched_schedule: 2620.967897310 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 5870
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967900112	5722	mm.page_alloc: 2620.967900112 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3	kernel	1	page_fault_entry	2620	967900213	5870	kernel.page_fault_entry: 2620.967900213 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 587
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_exit	2620	967910563	5722	kernel.page_fault_exit: 2620.967910563 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_entry	2620	967913002	5722	kernel.page_fault_entry: 2620.967913002 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3	mm	1	page_alloc	2620	967913103	5870	mm.page_alloc: 2620.967913103 (/home/frowand/tmp_ltt/trace3/mm_1), 5870, 5870, bash, , 5
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967914796	5722	mm.page_alloc: 2620.967914796 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_exit	2620	967920499	5722	kernel.page_fault_exit: 2620.967920499 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722,
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_entry	2620	967922911	5722	kernel.page_fault_entry: 2620.967922911 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967925012	5722	mm.page_alloc: 2620.967925012 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5

Time Frame start: 2620 s 967877373 ns end: 2620 s 967907889 ns Time Interval: 0 s 30516 ns Current Time: 2620 s 967900112 ns

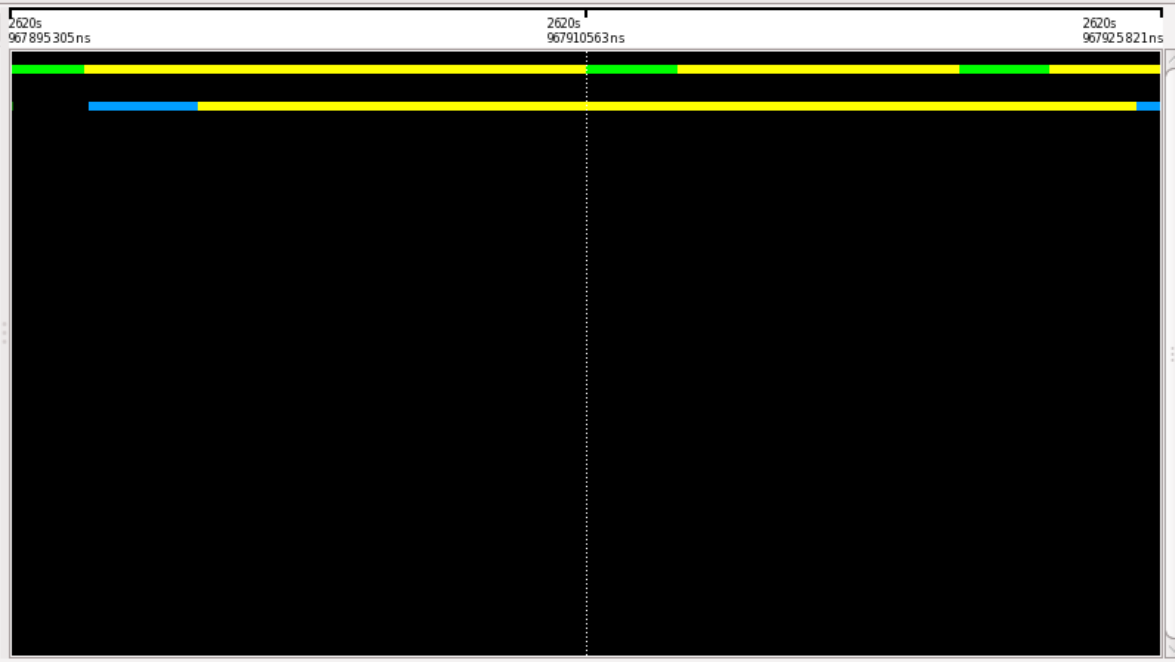

Linux Trace Toolkit Viewer
File View Tools Plugins
Help
Traceset
Process Brand PID TGID PPID CPU Birth sec Birth nsec TRACE
bash 5722 5722 5718 0 2620 136733380 0
bash 5870 5870 5722 0 2620 967886529 0
2620s 967877373ns
2620s 967892631ns
2620s 967907889ns
Trace Tracefile CPUID Event Time (s) Time (ns) PID Event Description
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967868935 5722 mm.page_alloc: 2620.967868935 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967873311 5722 mm.page_alloc: 2620.967873311 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 0 process_fork 2620 967886529 5722 kernel.process_fork: 2620.967886529 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, ba
/home/frowand/tmp_ltt/trace3 kernel 0 sched_wakeup_new_task 2620 967891447 5722 kernel.sched_wakeup_new_task: 2620.967891447 (/home/frowand/tmp_ltt/trace3/kernel_0), 5
/home/frowand/tmp_ltt/trace3 kernel 0 syscall_exit 2620 967894441 5722 kernel.syscall_exit: 2620.967894441 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, bas
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967897210 5722 kernel.page_fault_entry: 2620.967897210 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 kernel 1 sched_schedule 2620 967897310 5870 kernel.sched_schedule: 2620.967897310 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 5870
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967900112 5722 mm.page_alloc: 2620.967900112 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 1 page_fault_entry 2620 967900213 5870 kernel.page_fault_entry: 2620.967900213 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 587
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_exit 2620 967910563 5722 kernel.page_fault_exit: 2620.967910563 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967913002 5722 kernel.page_fault_entry: 2620.967913002 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 mm 1 page_alloc 2620 967913103 5870 mm.page_alloc: 2620.967913103 (/home/frowand/tmp_ltt/trace3/mm_1), 5870, 5870, bash, , 5
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967914796 5722 mm.page_alloc: 2620.967914796 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_exit 2620 967920499 5722 kernel.page_fault_exit: 2620.967920499 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722,
/home/frowand/tmp_ltt/trace3 kernel 0 page_fault_entry 2620 967922911 5722 kernel.page_fault_entry: 2620.967922911 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3 mm 0 page_alloc 2620 967925012 5722 mm.page_alloc: 2620.967925012 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
Time Frame start: 2620 s 967877373 ns end: 2620 s 967907889 ns Time Interval: 0 s 30516 ns
Current Time: 2620 s 967900213 ns

Linux Trace Toolkit Viewer

File View Tools Plugins Help

Traceset

Process	Brand	PID	TGID	PPID	CPU	Birth sec	Birth nsec	TRACE
bash		5722	5722	5718	0	2620	136733380	0
bash		5870	5870	5722	0	2620	967886529	0



Trace	Tracefile	CPUID	Event	Time (s)	Time (ns)	PID	Event Description
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967868935	5722	mm.page_alloc: 2620.967868935 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967873311	5722	mm.page_alloc: 2620.967873311 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3	kernel	0	process_fork	2620	967886529	5722	kernel.process_fork: 2620.967886529 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, ba
/home/frowand/tmp_ltt/trace3	kernel	0	sched_wakeup_new_task	2620	967891447	5722	kernel.sched_wakeup_new_task: 2620.967891447 (/home/frowand/tmp_ltt/trace3/kernel_0), 5
/home/frowand/tmp_ltt/trace3	kernel	0	syscall_exit	2620	967894441	5722	kernel.syscall_exit: 2620.967894441 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722, bas
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_entry	2620	967897210	5722	kernel.page_fault_entry: 2620.967897210 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3	kernel	1	sched_schedule	2620	967897310	5870	kernel.sched_schedule: 2620.967897310 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 5870
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967900112	5722	mm.page_alloc: 2620.967900112 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3	kernel	1	page_fault_entry	2620	967900213	5870	kernel.page_fault_entry: 2620.967900213 (/home/frowand/tmp_ltt/trace3/kernel_1), 5870, 587
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_exit	2620	967910563	5722	kernel.page_fault_exit: 2620.967910563 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_entry	2620	967913002	5722	kernel.page_fault_entry: 2620.967913002 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3	mm	1	page_alloc	2620	967913103	5870	mm.page_alloc: 2620.967913103 (/home/frowand/tmp_ltt/trace3/mm_1), 5870, 5870, bash, , 5
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967914796	5722	mm.page_alloc: 2620.967914796 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_exit	2620	967920499	5722	kernel.page_fault_exit: 2620.967920499 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 5722,
/home/frowand/tmp_ltt/trace3	kernel	0	page_fault_entry	2620	967922911	5722	kernel.page_fault_entry: 2620.967922911 (/home/frowand/tmp_ltt/trace3/kernel_0), 5722, 572
/home/frowand/tmp_ltt/trace3	mm	0	page_alloc	2620	967925012	5722	mm.page_alloc: 2620.967925012 (/home/frowand/tmp_ltt/trace3/mm_0), 5722, 5722, bash, , 5

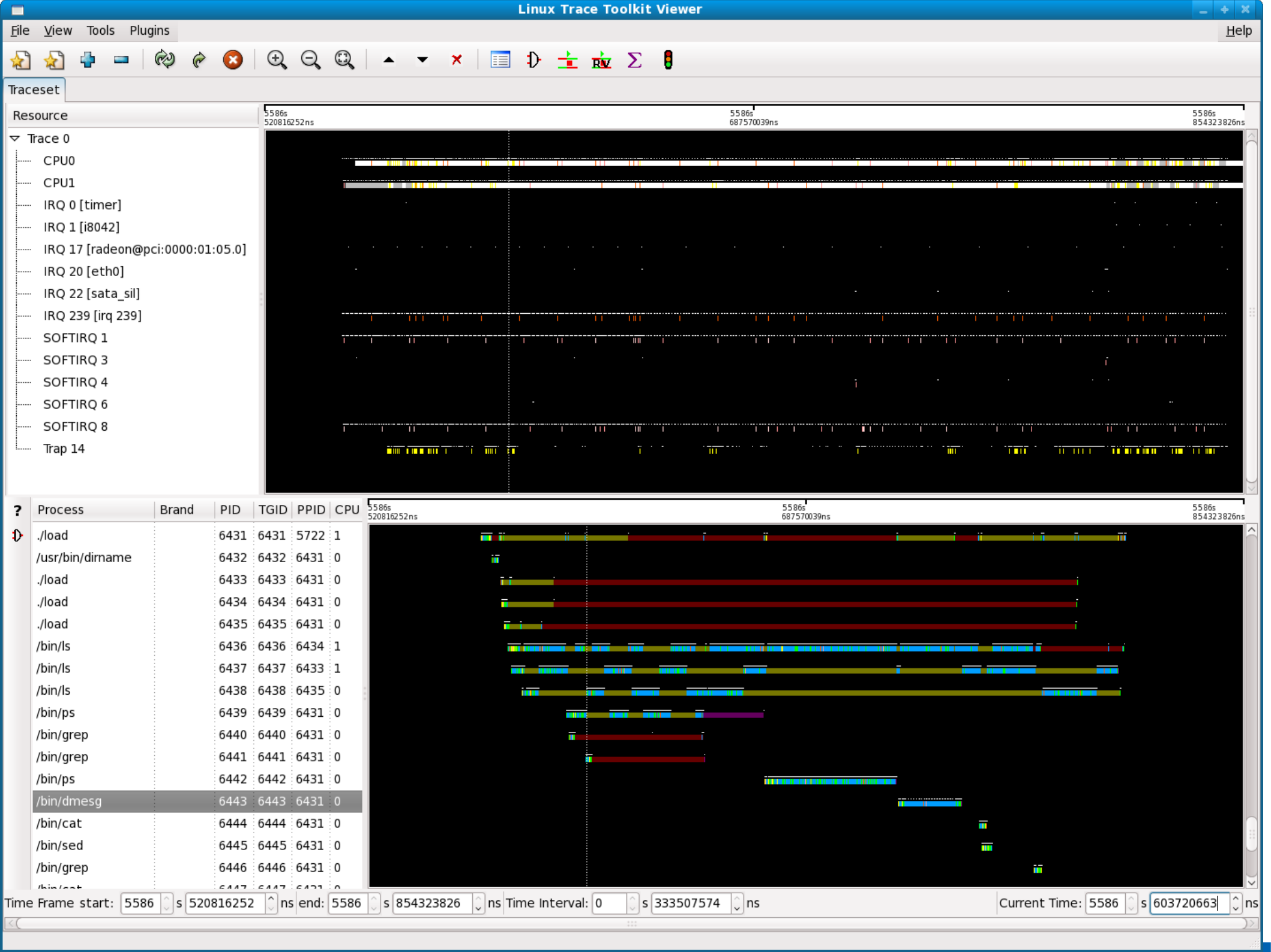
Time Frame start: 2620 s 967895305 ns end: 2620 s 967925821 ns Time Interval: 0 s 30516 ns Current Time: 2620 s 967910563 ns

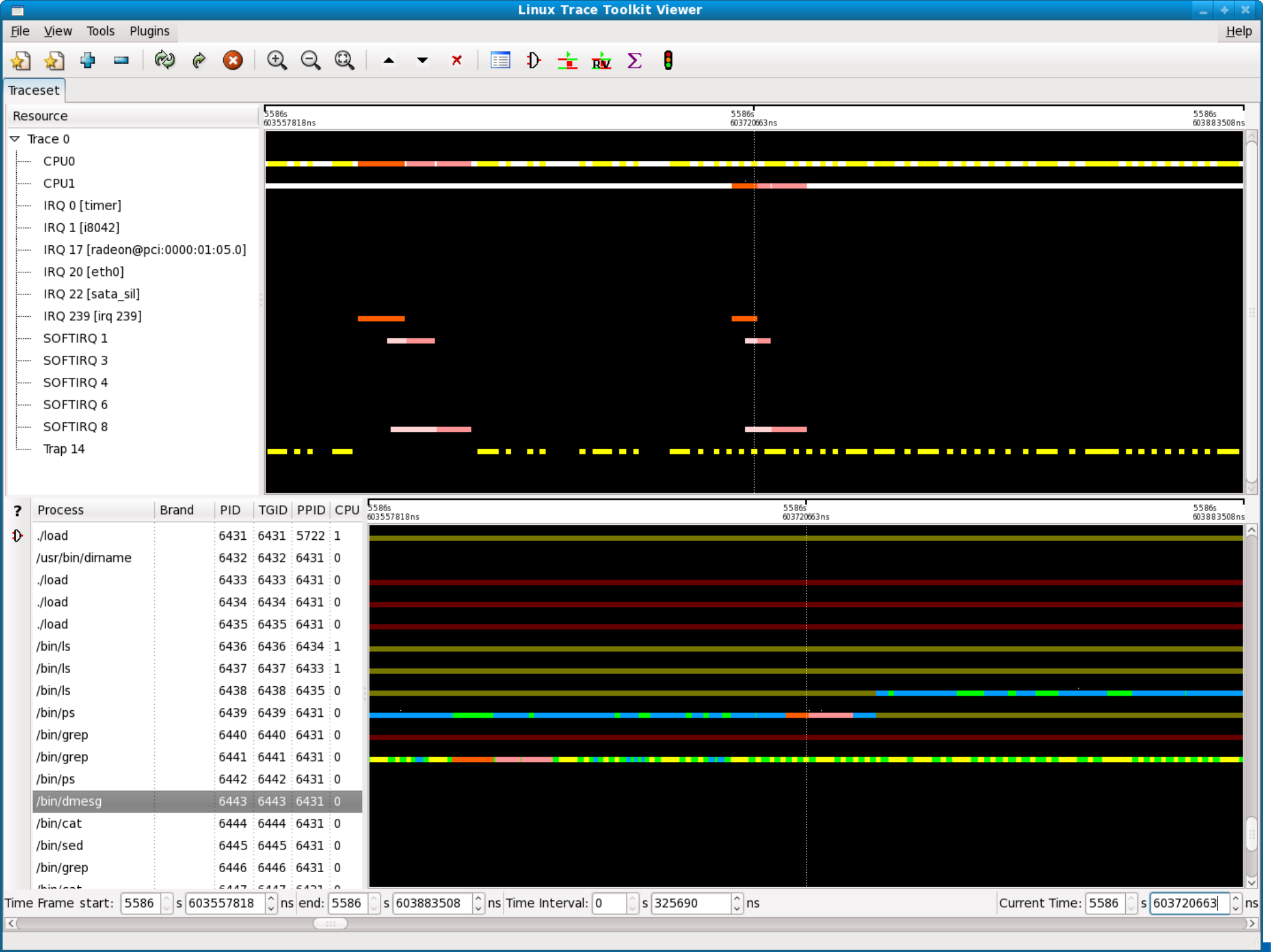
lttng [timeline]

Text dump of logfile, reformatted for my taste:

```
lttv -m textDump -e "state.pid=5722|state.pid=5870" -t trace3 >tmp

trace_path=/mnt/fc8/frowand_me/presentation/
survey_of_linux_measurement_and_diagnostic_tools/
preparation/resources/ltt/trace3/

sed -i -e "s|${trace_path}||" tmp

paste <(cut -d" " -f 2 tmp) <(cut -d" " -f 1 tmp) <(cut -d" " -f 3- tmp) >td_5722_5870
```

```
2620.967632062  kernel.page_fault_entry:        (kernel_0), 5722, 5722, bash, , 5718, 0x0, TRAP { ip = 0x44e6d8, address = 0x8ff3dc, trap_id = 14, write_access = 1 }
2620.967635129  kernel.page_fault_exit: (kernel_0), 5722, 5722, bash, , 5718, 0x0, USER_MODE { res = 8 }
2620.967666667  kernel.page_fault_entry:        (kernel_0), 5722, 5722, bash, , 5718, 0x0, TRAP { ip = 0x44e6d8, address = 0x8e2dfc, trap_id = 14, write_access = 1 }
2620.967668332  kernel.page_fault_exit: (kernel_0), 5722, 5722, bash, , 5718, 0x0, USER_MODE { res = 8 }
2620.967674811  kernel.page_fault_entry:        (kernel_0), 5722, 5722, bash, , 5718, 0x0, TRAP { ip = 0x4338ca, address = 0x904d68, trap_id = 14, write_access = 1 }
2620.967676271  kernel.page_fault_exit: (kernel_0), 5722, 5722, bash, , 5718, 0x0, USER_MODE { res = 8 }
2620.967679023  kernel.syscall_entry:    (kernel_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { ip = 0x361b0335b0, syscall_id = 14 [sys_rt_sigprocmask+0x0/0xdc] }
2620.967680477  kernel.syscall_exit:     (kernel_0), 5722, 5722, bash, , 5718, 0x0, USER_MODE { ret = 0 }
2620.967682150  kernel.syscall_entry:    (kernel_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { ip = 0x361b0d1b57, syscall_id = 22 [sys_pipe+0x0/0x12] }
2620.967694120  kernel.syscall_exit:     (kernel_0), 5722, 5722, bash, , 5718, 0x0, USER_MODE { ret = 0 }
2620.967695994  kernel.syscall_entry:    (kernel_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { ip = 0x361b0a4096, syscall_id = 56 [stub_clone+0x0/0x20] }
2620.967708325  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 411892, order = 1 }
2620.967726416  kernel.sched_migrate_task:       (kernel_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pid = 5722, state = 0, dest_cpu = 1 }
2620.967743016  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 421256, order = 0 }
2620.967757693  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 411921, order = 0 }
2620.967761232  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 454864, order = 0 }
2620.967764552  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 494885, order = 0 }
2620.967778499  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 411841, order = 0 }
2620.967794566  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 454875, order = 0 }
2620.967797364  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 421068, order = 0 }
2620.967809370  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 411863, order = 0 }
2620.967823816  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 411911, order = 0 }
2620.967834235  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 411702, order = 0 }
2620.967842895  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 411821, order = 0 }
2620.967845842  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 411848, order = 0 }
2620.967848821  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 411859, order = 0 }
2620.967859644  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 420979, order = 0 }
2620.967866031  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 411856, order = 0 }
2620.967868935  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 421158, order = 0 }
2620.967873311  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pfn = 392391, order = 0 }

2620.967886529  kernel.process_fork:     (kernel_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { parent_pid = 5722, child_pid = 5870, child_tgid = 5870 }
2620.967891447  kernel.sched_wakeup_new_task:    (kernel_0), 5722, 5722, bash, , 5718, 0x0, SYSCALL { pid = 5870, state = 0, cpu_id = 1 }
2620.967894441  kernel.syscall_exit:     (kernel_0), 5722, 5722, bash, , 5718, 0x0, USER_MODE { ret = 5870 }
2620.967897210  kernel.page_fault_entry:        (kernel_0), 5722, 5722, bash, , 5718, 0x0, TRAP { ip = 0x361b0a41f5, address = 0x7f4a27feb7c4, trap_id = 14, write_access = 1 }
2620.967897310  kernel.sched_schedule:  (kernel_1), 5870, 5870, bash, , 5722, 0x0, SYSCALL { prev_pid = 0, next_pid = 5870, prev_state = 0 }
2620.967900112  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, TRAP { pfn = 422927, order = 0 }
2620.967900213  kernel.page_fault_entry:        (kernel_1), 5870, 5870, bash, , 5722, 0x0, TRAP { ip = 0xffffffff811d1dbd, address = 0x7f4a27feb7c0, trap_id = 14, write_access = 1 }
2620.967910563  kernel.page_fault_exit: (kernel_0), 5722, 5722, bash, , 5718, 0x0, USER_MODE { res = 8 }


2620.967913002  kernel.page_fault_entry:        (kernel_0), 5722, 5722, bash, , 5718, 0x0, TRAP { ip = 0x361b0a41fe, address = 0x7fffadfa8ec8, trap_id = 14, write_access = 1 }
2620.967913103  mm.page_alloc:   (mm_1), 5870, 5870, bash, , 5722, 0x0, TRAP { pfn = 450925, order = 0 }
2620.967914796  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, TRAP { pfn = 422926, order = 0 }
2620.967920499  kernel.page_fault_exit: (kernel_0), 5722, 5722, bash, , 5718, 0x0, USER_MODE { res = 8 }
2620.967922911  kernel.page_fault_entry:        (kernel_0), 5722, 5722, bash, , 5718, 0x0, TRAP { ip = 0x361b073eab, address = 0x361b369dd4, trap_id = 14, write_access = 1 }
2620.967925012  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, TRAP { pfn = 422360, order = 0 }
2620.967925112  mm.page_free:    (mm_1), 5870, 5870, bash, , 5722, 0x0, TRAP { pfn = 414270, order = 0 }
2620.967925213  kernel.page_fault_exit: (kernel_1), 5870, 5870, bash, , 5722, 0x0, SYSCALL { res = 8 }
2620.967930466  kernel.page_fault_exit: (kernel_0), 5722, 5722, bash, , 5718, 0x0, USER_MODE { res = 8 }
2620.967930567  kernel.syscall_exit:     (kernel_1), 5870, 5870, bash, , 5722, 0x0, USER_MODE { ret = 0 }
2620.967933149  kernel.page_fault_entry:        (kernel_0), 5722, 5722, bash, , 5718, 0x0, TRAP { ip = 0x361b075459, address = 0x361b36a7f0, trap_id = 14, write_access = 1 }
2620.967935023  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, TRAP { pfn = 419977, order = 0 }
2620.967935124  kernel.page_fault_entry:        (kernel_1), 5870, 5870, bash, , 5722, 0x0, TRAP { ip = 0x361b0a4096, address = 0x361b0a4096, trap_id = 14, write_access = 0 }
2620.967935224  mm.page_alloc:   (mm_1), 5870, 5870, bash, , 5722, 0x0, TRAP { pfn = 392382, order = 0 }
2620.967940110  kernel.page_fault_exit: (kernel_0), 5722, 5722, bash, , 5718, 0x0, USER_MODE { res = 8 }
2620.967942593  kernel.page_fault_entry:        (kernel_0), 5722, 5722, bash, , 5718, 0x0, TRAP { ip = 0x361b07548c, address = 0x361b3684f8, trap_id = 14, write_access = 1 }
2620.967944441  mm.page_alloc:   (mm_0), 5722, 5722, bash, , 5718, 0x0, TRAP { pfn = 419881, order = 0 }
2620.967944542  kernel.page_fault_exit: (kernel_1), 5870, 5870, bash, , 5722, 0x0, USER_MODE { res = 512 }
```

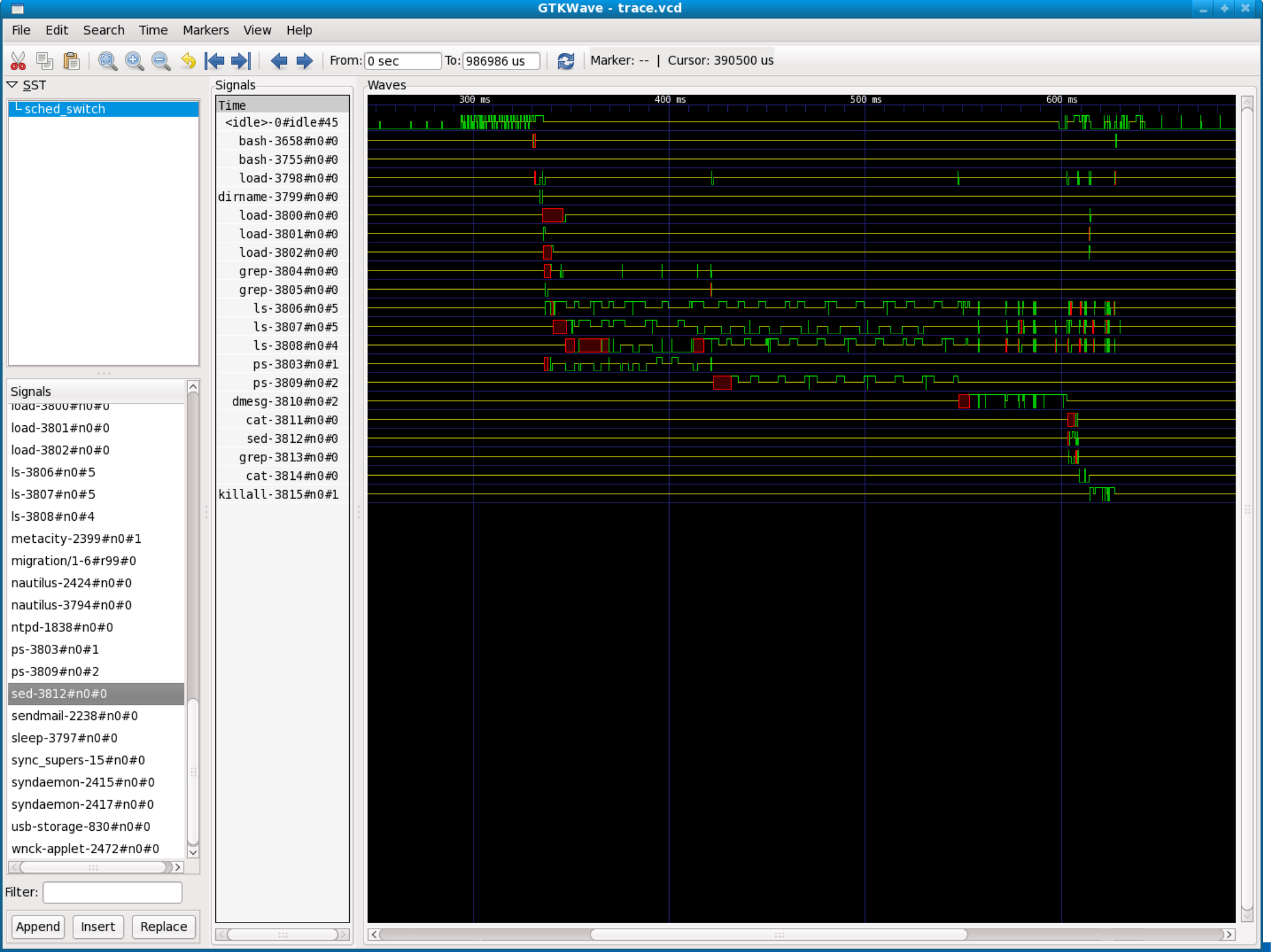

Linux Trace Toolkit Viewer
File View Tools Plugins Help
Traceset
Resource
Trace 0
CPU0
CPU1
IRQ 0 [timer]
IRQ 1 [i8042]
IRQ 17 [radeon@pci:0000:01:05.0]
IRQ 20 [eth0]
IRQ 22 [sata_sil]
IRQ 239 [irq 239]
SOFTIRQ 1
SOFTIRQ 3
SOFTIRQ 4
SOFTIRQ 6
SOFTIRQ 8
Trap 14
5586s 520816252ns
5586s 687570039ns
5586s 854323826ns
Process Brand PID TGID PPID CPU
./load 6431 6431 5722 1
/usr/bin/dirname 6432 6432 6431 0
./load 6433 6433 6431 0
./load 6434 6434 6431 0
./load 6435 6435 6431 0
/bin/ls 6436 6436 6434 1
/bin/ls 6437 6437 6433 1
/bin/ls 6438 6438 6435 0
/bin/ps 6439 6439 6431 0
/bin/grep 6440 6440 6431 0
/bin/grep 6441 6441 6431 0
/bin/ps 6442 6442 6431 0
/bin/dmesg 6443 6443 6431 0
/bin/cat 6444 6444 6431 0
/bin/sed 6445 6445 6431 0
/bin/grep 6446 6446 6431 0
Time Frame start: 5586 s 520816252 ns end: 5586 s 854323826 ns Time Interval: 0 s 333507574 ns Current Time: 5586 s 603720663 ns

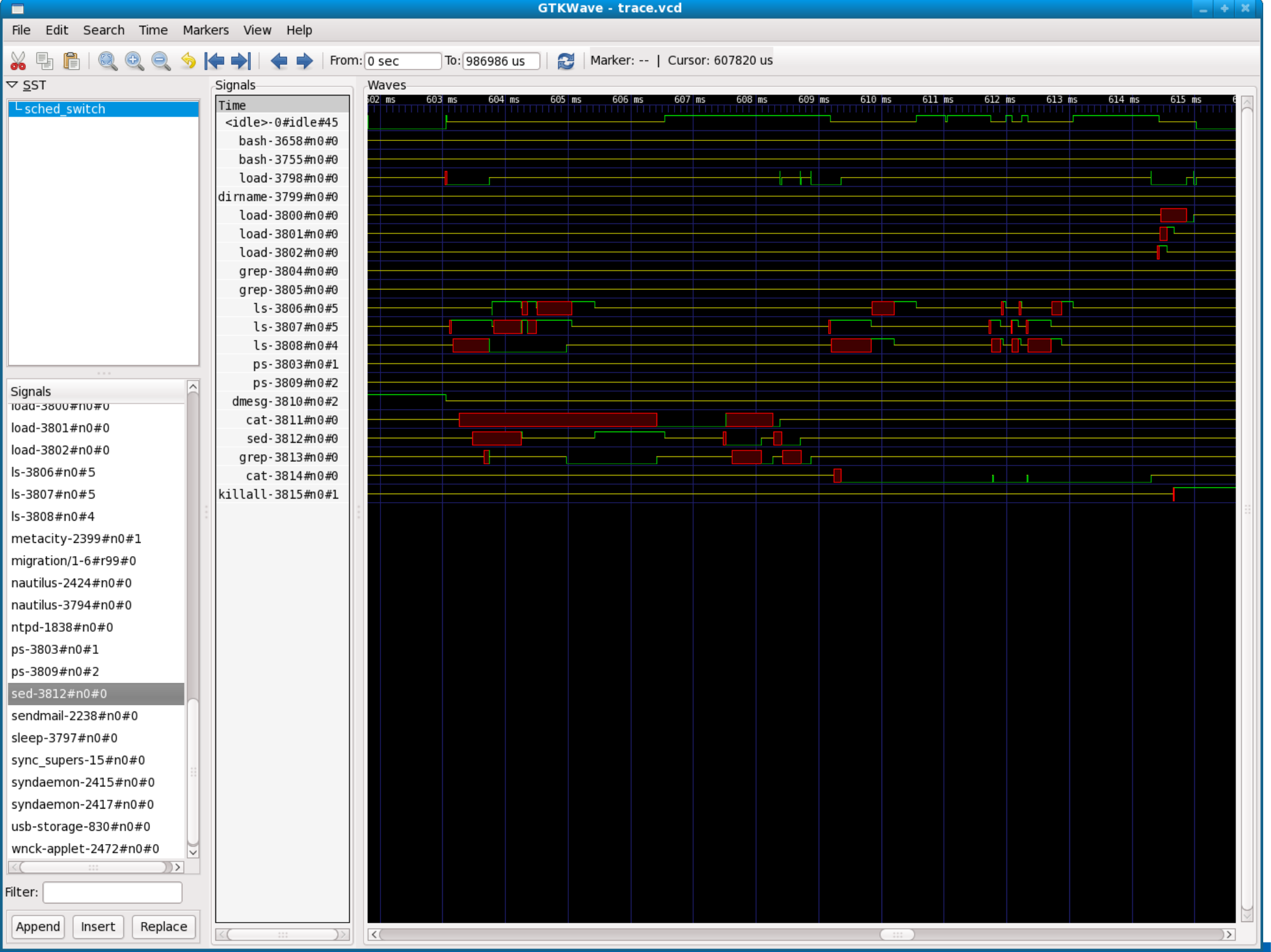

Linux Trace Toolkit Viewer
File View Tools Plugins Help
Traceset
Resource
Trace 0
CPU0
CPU1
IRQ 0 [timer]
IRQ 1 [i8042]
IRQ 17 [radeon@pci:0000:01:05.0]
IRQ 20 [eth0]
IRQ 22 [sata_sil]
IRQ 239 [irq 239]
SOFTIRQ 1
SOFTIRQ 3
SOFTIRQ 4
SOFTIRQ 6
SOFTIRQ 8
Trap 14
5586s 603557818ns
5586s 603720663ns
5586s 603883508ns
Process Brand PID TGID PPID CPU
./load 6431 6431 5722 1
/usr/bin/dirname 6432 6432 6431 0
./load 6433 6433 6431 0
./load 6434 6434 6431 0
./load 6435 6435 6431 0
/bin/ls 6436 6436 6434 1
/bin/ls 6437 6437 6433 1
/bin/ls 6438 6438 6435 0
/bin/ps 6439 6439 6431 0
/bin/grep 6440 6440 6431 0
/bin/grep 6441 6441 6431 0
/bin/ps 6442 6442 6431 0
/bin/dmesg 6443 6443 6431 0
/bin/cat 6444 6444 6431 0
/bin/sed 6445 6445 6431 0
/bin/grep 6446 6446 6431 0
Time Frame start: 5586 s 603557818 ns end: 5586 s 603883508 ns Time Interval: 0 s 325690 ns Current Time: 5586 s 603720663 ns

sched_switch [timeline]

Graphic representation of ftrace sched_switch tracer.

The following graphs show the same toy benchmark as in the last two ltt graphs.

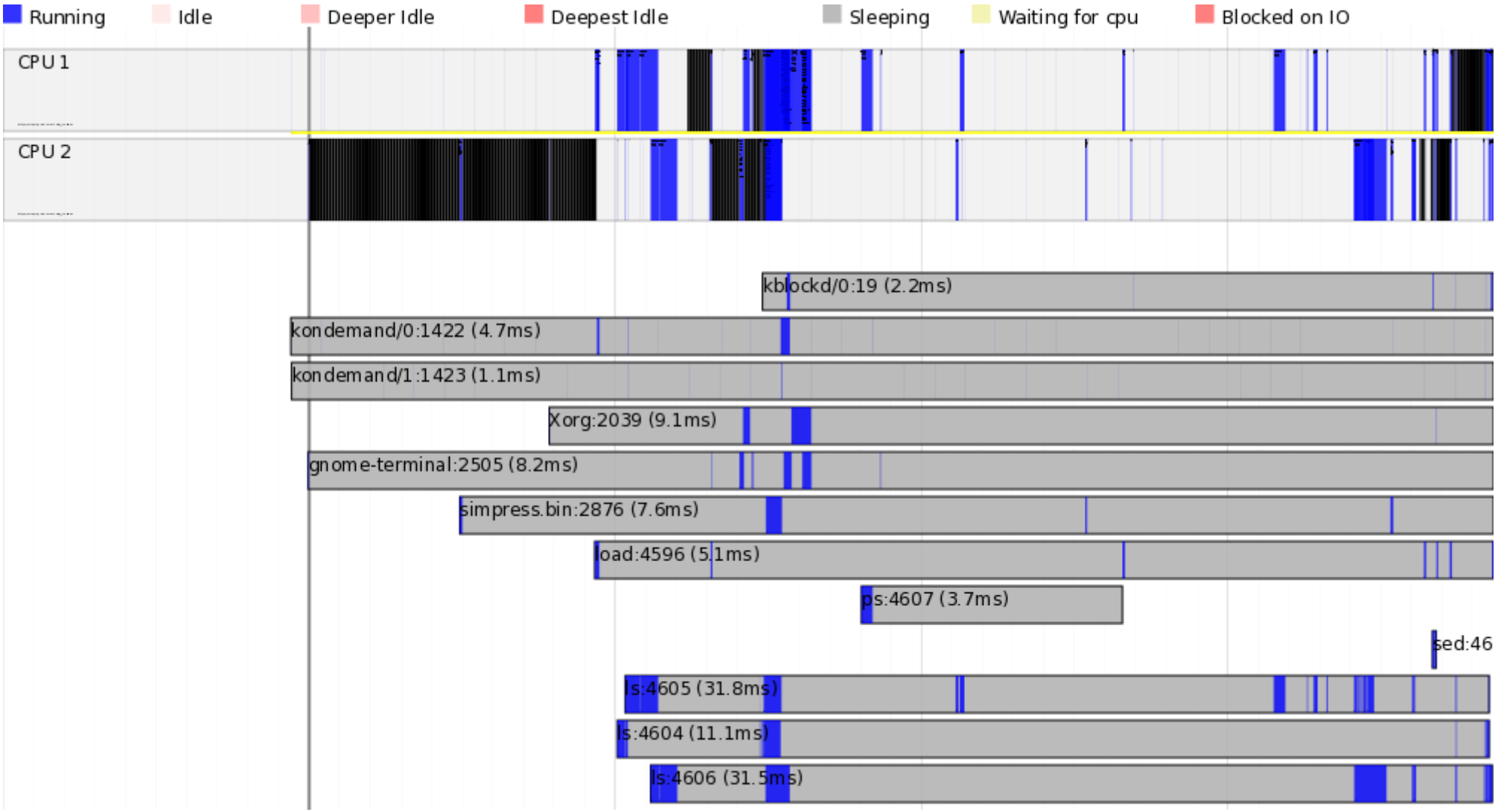

GTKWave - trace.vcd
File
Edit
Search
Time
Markers
View
Help
From: 0 sec
To: 986986 us
Marker: -- | Cursor: 390500 us
SST
sched_switch
Signals
load-3801#n0#0
load-3802#n0#0
ls-3806#n0#5
ls-3807#n0#5
ls-3808#n0#4
metacity-2399#n0#1
migration/1-6#r99#0
nautilus-2424#n0#0
nautilus-3794#n0#0
ntpd-1838#n0#0
ps-3803#n0#1
ps-3809#n0#2
sed-3812#n0#0
sendmail-2238#n0#0
sleep-3797#n0#0
sync_supers-15#n0#0
syndaemon-2415#n0#0
syndaemon-2417#n0#0
usb-storage-830#n0#0
wnck-applet-2472#n0#0
Filter:
Append
Insert
Replace
Signals
Time
<idle>-0#idle#45
bash-3658#n0#0
bash-3755#n0#0
load-3798#n0#0
dirname-3799#n0#0
load-3800#n0#0
load-3801#n0#0
load-3802#n0#0
grep-3804#n0#0
grep-3805#n0#0
ls-3806#n0#5
ls-3807#n0#5
ls-3808#n0#4
ps-3803#n0#1
ps-3809#n0#2
dmesg-3810#n0#2
cat-3811#n0#0
sed-3812#n0#0
grep-3813#n0#0
cat-3814#n0#0
killall-3815#n0#1
Waves
300 ms
400 ms
500 ms
600 ms


GTKWave - trace.vcd
File Edit Search Time Markers View Help
From: 0 sec
To: 986986 us
Marker: -- | Cursor: 607820 us
SST
sched_switch
Signals
load-3801#n0#0
load-3802#n0#0
ls-3806#n0#5
ls-3807#n0#5
ls-3808#n0#4
metacity-2399#n0#1
migration/1-6#r99#0
nautilus-2424#n0#0
nautilus-3794#n0#0
ntpd-1838#n0#0
ps-3803#n0#1
ps-3809#n0#2
sed-3812#n0#0
sendmail-2238#n0#0
sleep-3797#n0#0
sync_supers-15#n0#0
syndaemon-2415#n0#0
syndaemon-2417#n0#0
usb-storage-830#n0#0
wnck-applet-2472#n0#0
Filter:
Append
Insert
Replace
Signals
Time
<idle>-0#idle#45
bash-3658#n0#0
bash-3755#n0#0
load-3798#n0#0
dirname-3799#n0#0
load-3800#n0#0
load-3801#n0#0
load-3802#n0#0
grep-3804#n0#0
grep-3805#n0#0
ls-3806#n0#5
ls-3807#n0#5
ls-3808#n0#4
ps-3803#n0#1
ps-3809#n0#2
dmesg-3810#n0#2
cat-3811#n0#0
sed-3812#n0#0
grep-3813#n0#0
cat-3814#n0#0
killall-3815#n0#1
Waves
603 ms
604 ms
605 ms
606 ms
607 ms
608 ms
609 ms
610 ms
611 ms
612 ms
613 ms
614 ms
615 ms

timechart [timeline]

Graphic representation of performance events data.

The following graph shows the same toy benchmark as in the last two ltt graphs and the sched_switch graphs.


Running
Idle
Deeper Idle
Deepest Idle
Sleeping
Waiting for cpu
Blocked on IO
CPU 1
CPU 2
kblockd/0:19 (2.2ms)
kondemand/0:1422 (4.7ms)
kondemand/1:1423 (1.1ms)
Xorg:2039 (9.1ms)
gnome-terminal:2505 (8.2ms)
simpress.bin:2876 (7.6ms)
load:4596 (5.1ms)
ps:4607 (3.7ms)
sed:46
ls:4605 (31.8ms)
ls:4604 (11.1ms)
ls:4606 (31.5ms)

ftrace

ftrace is a key set of performance tools and instrumentation.

Not welll covered in this talk because
- it is a huge topic, deserving its own talk
- information is otherwise available

In mainline 2.6.27
Rapidly evolving

Performance Events

Also known as:

- Performance Counters
- perfcounters

Performance Events

Performance Events is a key set of performance tools and instrumentation.

Not well covered in this talk because
- it is a huge topic, deserving its own talk
- information is otherwise available

Rapidly evolving

Past vs. Future

lkml: Re: mailing list for trace users
Ingo Molnar
Wed, 23 Sep 2009 22:07:47 +0200

> “Both oprofile and readprof are obsolete.”

http://dev.linuxfoundation.org/~tytso/tab-minutes/2009-01.html

> “Maybe we need to socialize in April the fact that Systemtap is doomed, and to socialize exploring other tracing technologies.

Past vs. Future

Are predictions of obsolescence correct?

Review

Some past, current and future tools to investigate:

- Waiting for resources
- Resource usage
- Resource allocation

Resources

latencytop
  http://www.latencytop.org/

mutrace
  http://0pointer.de/blog/projects/mutrace.html

smem
  http://www.selenic.com/smem/
  http://tree.celinuxforum.org/CelfPubWiki/ELC2009Presentations
    Matt Mackall "Visualizing Process Memory"
  http://selenic.com/mailman/listinfo/smem

systemtap
  http://sourceware.org/systemtap/
  /usr/share/doc/system-tap*/examples/

Resources

bootchart
  http://www.bootchart.org/

kft
  http://elinux.org/Kernel_Function_Trace
  http://elinux.org/Using_Kernel_Function_Trace

Analyzing Kernel Function Execution with ftrace
  http://embeddedlinuxconference.com/elc_europe09/
  http://embeddedlinuxconference.com/elc_europe09/sessions.html#Bird

Resources

ltt ng
  http://lttng.org/

sched_switch
  linux-rt-users (http://vger.kernel.org)
  Analyze sched_switch ftrace data with vcd viewer
  Herman ten Brugge <hermantenbrugge@xxxxxxx>
  Thu, 04 Jun 2009 20:58:57 +0200

timechart
  http://blog.fenrus.org/?p=5

Questions?