KernelShark
(quick tutorial)

Steven Rostedt
srostedt@redhat.com
rostedt@goodmis.org

trace-cmd

- binary tool to read Ftrace's buffers
  - Records into a trace.dat file for later reads
  - Reads the trace.dat file
    - Can record on big endian, read in little, and vice versa
  - Reads the raw buffers using splice
  - Will automatically mount debugfs if it is not mounted

trace-cmd record

- Default, writes to “trace.dat”

```
[root@frodo ~]# trace-cmd record -e sched ls -ltr /usr > /dev/null
disable all
enable sched
offset=2f2000
offset=2f4000
[root@frodo ~]# trace-cmd record -o func.dat -p function ls -ltr /usr > /dev/null
  plugin function
disable all
offset=2f2000
offset=412000
[root@frodo ~]# trace-cmd record -o fgraph.dat -p function_graph ls -ltr /usr \
   > /dev/null
  plugin function_graph
disable all
offset=2f2000
offset=460000
[root@frodo ~]# trace-cmd record -o fgraph-events.dat -e sched -p function_graph \
    ls -ltr /usr > /dev/null
  plugin function_graph
disable all
enable sched
offset=2f2000
offset=461000
```

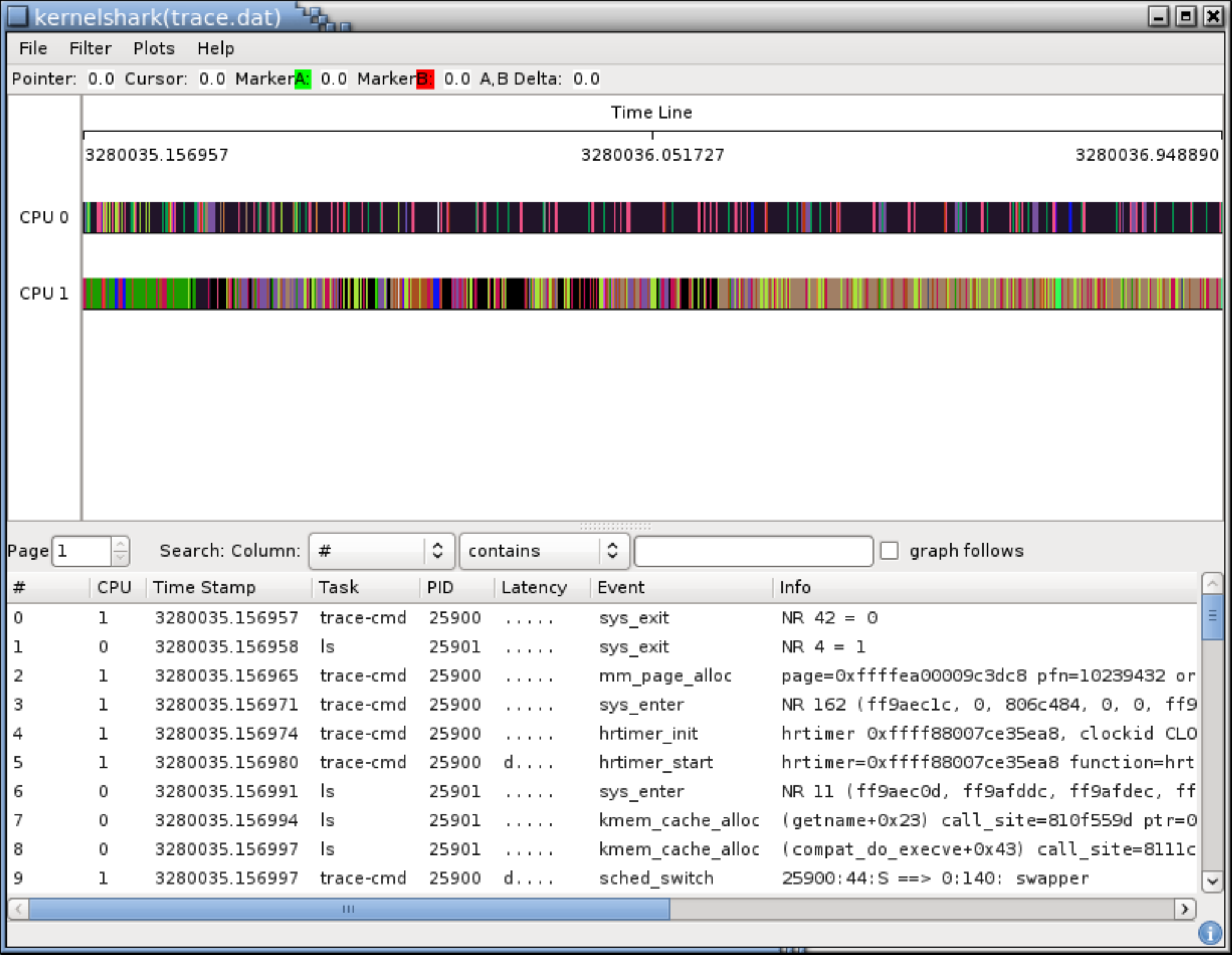
trace-cmd report

- Default, reads from “trace.dat”

```
[root@frodo ~]# trace-cmd report | head -15
version = 6
cpus=2
       trace-cmd-6157  [000]    83.713584: sched_stat_runtime:   task: trace-cmd:61
       trace-cmd-6157  [000]    83.713591: sched_switch:         6157:120:S ==> 0:1
          <idle>-0     [000]    83.713646: sched_stat_wait:      task: trace-cmd:61
          <idle>-0     [000]    83.713648: sched_switch:         0:120:R ==> 6158:1
              ls-6158  [001]    83.713934: sched_wakeup:         6158:?:? +   5900:
              ls-6158  [001]    83.713935: sched_stat_runtime:   task: trace-cmd:61
              ls-6158  [001]    83.713937: sched_stat_runtime:   task: trace-cmd:61
              ls-6158  [001]    83.713938: sched_switch:         6158:120:R ==> 590
    migration/1-5900   [001]    83.713941: sched_stat_wait:      task: trace-cmd:61
    migration/1-5900   [001]    83.713942: sched_migrate_task:   task trace-cmd:615
    migration/1-5900   [001]    83.713947: sched_switch:         5900:0:S ==> 0:120
              ls-6158  [000]    83.714067: sched_stat_runtime:   task: ls:6158 runt
              ls-6158  [000]    83.714636: sched_stat_runtime:   task: ls:6158 runt
```

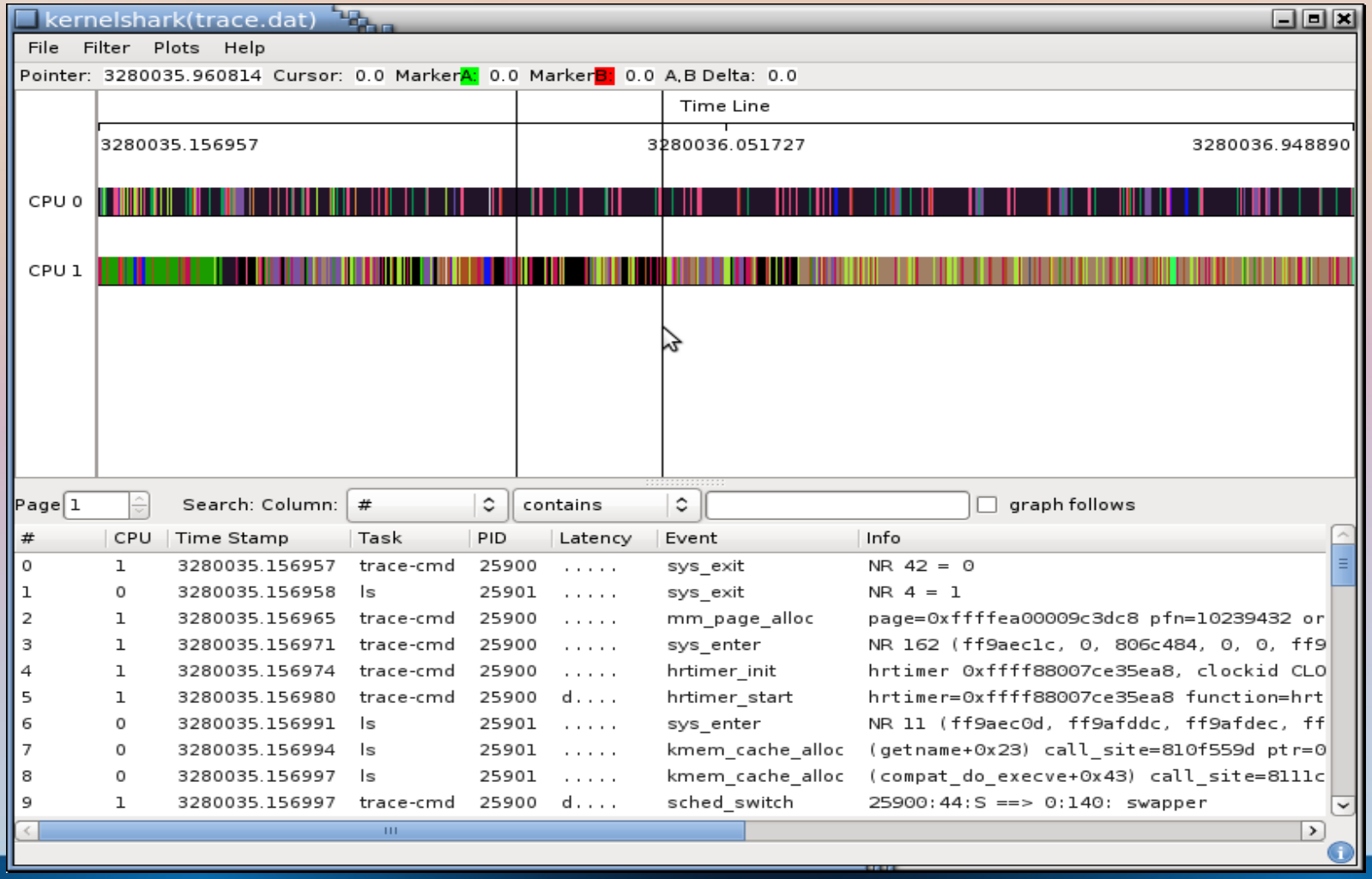
KernelShark

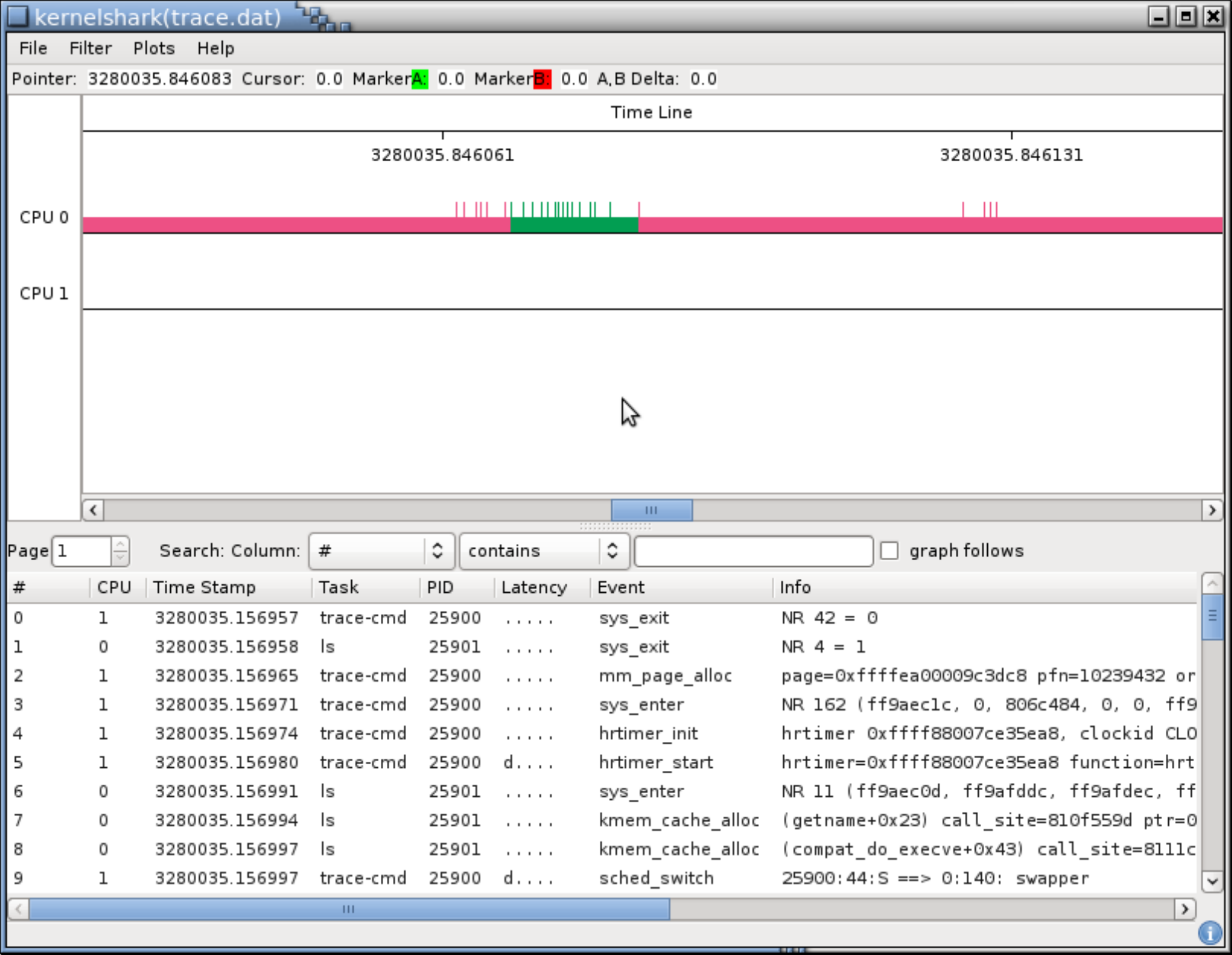
- A front end reader of the trace-cmd trace.dat file
- Graph view
- List view
- Simple and Advance filtering


kernelshark(trace.dat)
File
Filter
Plots
Help
Pointer: 0.0 Cursor: 0.0 MarkerA: 0.0 MarkerB: 0.0 A,B Delta: 0.0
Time Line
3280035.156957
3280036.051727
3280036.948890
CPU 0
CPU 1
Page 1
Search: Column:
#
contains
graph follows
CPU Time Stamp Task PID Latency Event Info
0 1 3280035.156957 trace-cmd 25900 ..... sys_exit NR 42 = 0
1 0 3280035.156958 ls 25901 ..... sys_exit NR 4 = 1
2 1 3280035.156965 trace-cmd 25900 ..... mm_page_alloc page=0xffffea00009c3dc8 pfn=10239432 or
3 1 3280035.156971 trace-cmd 25900 ..... sys_enter NR 162 (ff9aec1c, 0, 806c484, 0, 0, ff9
4 1 3280035.156974 trace-cmd 25900 ..... hrtimer_init hrtimer 0xffff88007ce35ea8, clockid CLO
5 1 3280035.156980 trace-cmd 25900 d.... hrtimer_start hrtimer=0xffff88007ce35ea8 function=hrt
6 0 3280035.156991 ls 25901 ..... sys_enter NR 11 (ff9aec0d, ff9afddc, ff9afdec, ff
7 0 3280035.156994 ls 25901 ..... kmem_cache_alloc (getname+0x23) call_site=810f559d ptr=0
8 0 3280035.156997 ls 25901 ..... kmem_cache_alloc (compat_do_execve+0x43) call_site=8111c
9 1 3280035.156997 trace-cmd 25900 d.... sched_switch 25900:44:S ==> 0:140: swapper

Zooming In

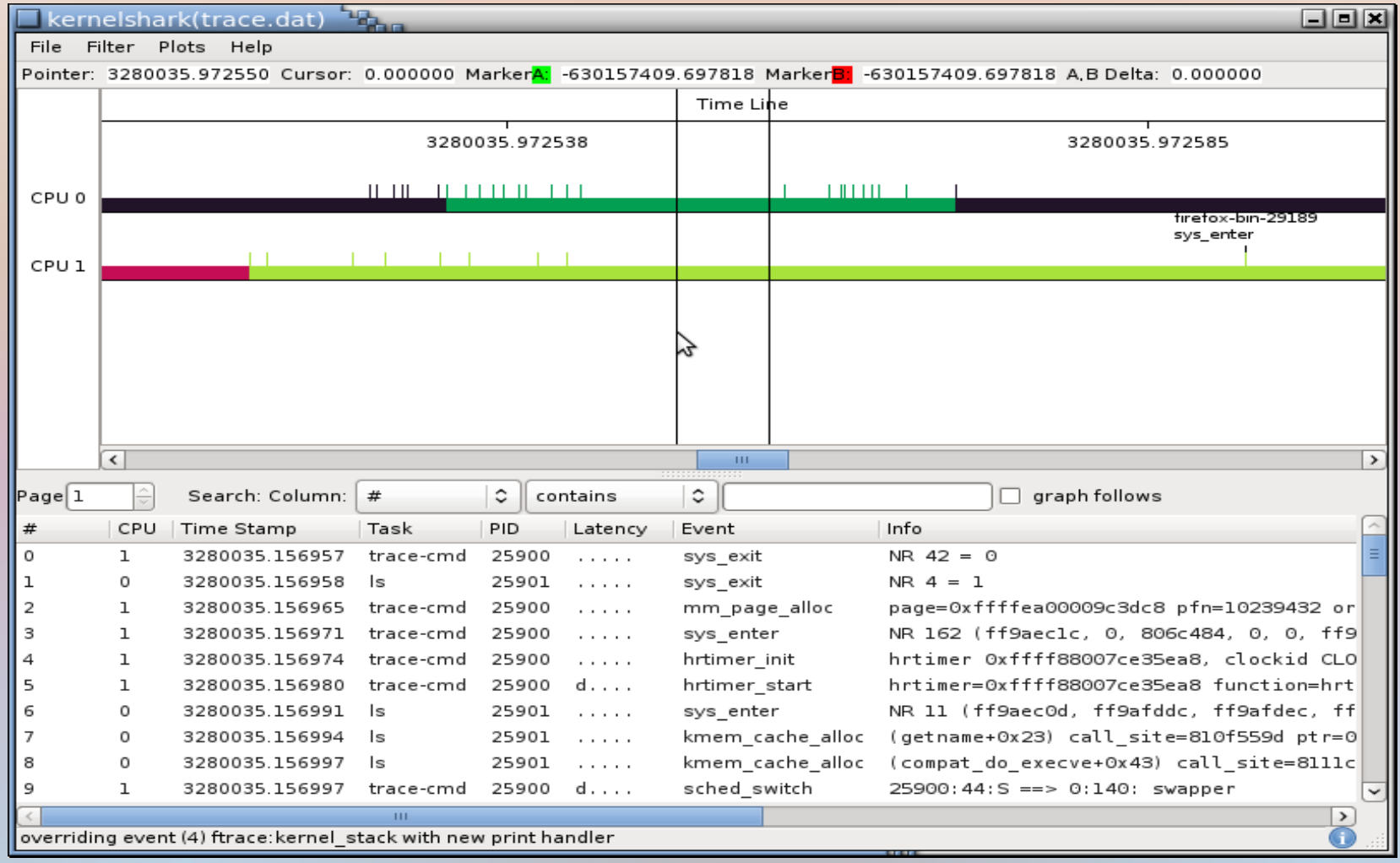
- Left click and drag to the right



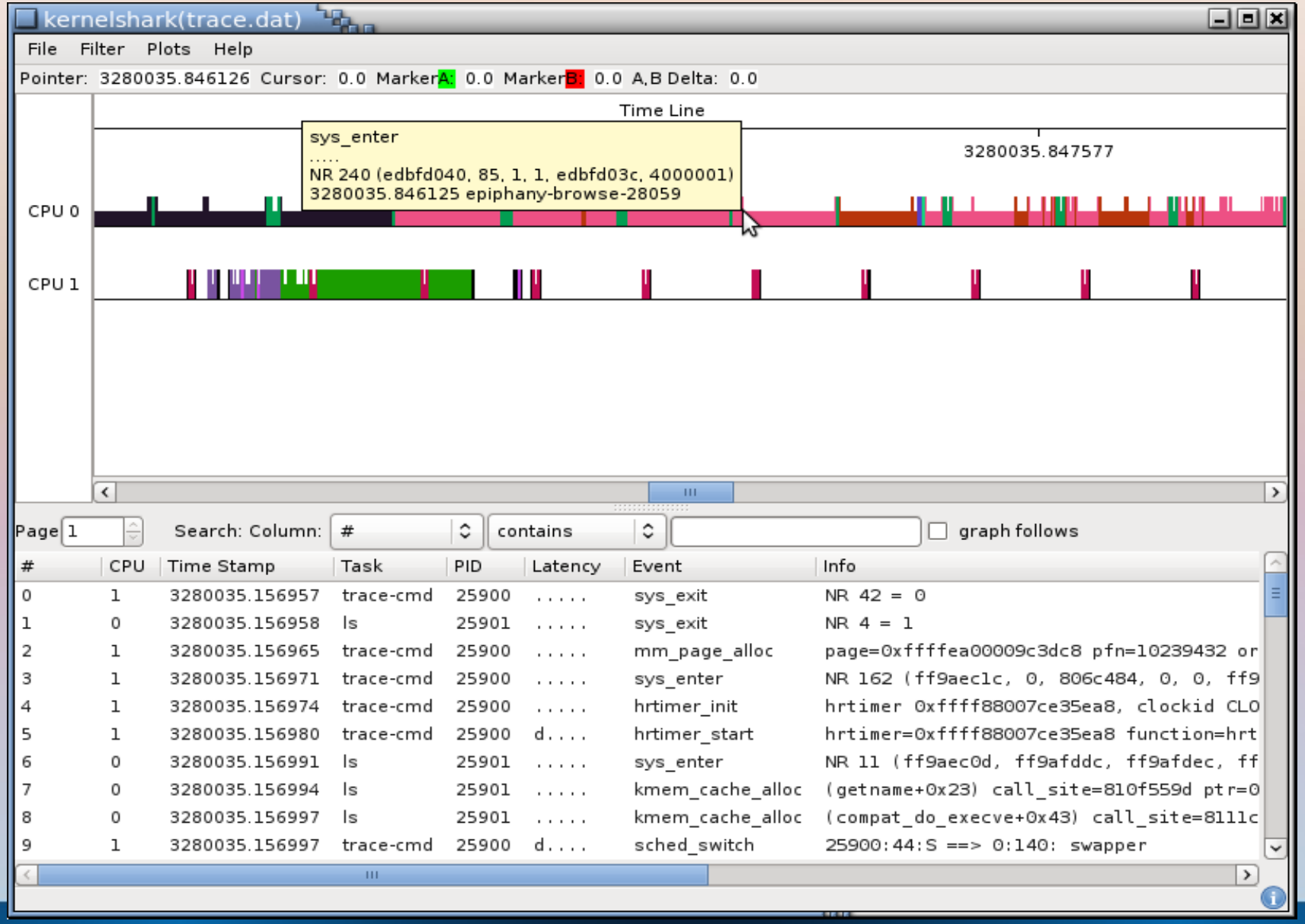

kernelshark(trace.dat)
File
Filter
Plots
Help
Pointer: 3280035.846083 Cursor: 0.0 MarkerA: 0.0 MarkerB: 0.0 A,B Delta: 0.0
Time Line
3280035.846061
3280035.846131
CPU 0
CPU 1
Page 1
Search: Column: #
contains
graph follows
CPU Time Stamp Task PID Latency Event Info
0 1 3280035.156957 trace-cmd 25900 ..... sys_exit NR 42 = 0
1 0 3280035.156958 ls 25901 ..... sys_exit NR 4 = 1
2 1 3280035.156965 trace-cmd 25900 ..... mm_page_alloc page=0xffffea00009c3dc8 pfn=10239432 or
3 1 3280035.156971 trace-cmd 25900 ..... sys_enter NR 162 (ff9aec1c, 0, 806c484, 0, 0, ff9
4 1 3280035.156974 trace-cmd 25900 ..... hrtimer_init hrtimer 0xffff88007ce35ea8, clockid CLO
5 1 3280035.156980 trace-cmd 25900 d.... hrtimer_start hrtimer=0xffff88007ce35ea8 function=hrt
6 0 3280035.156991 ls 25901 ..... sys_enter NR 11 (ff9aec0d, ff9afddc, ff9afdec, ff
7 0 3280035.156994 ls 25901 ..... kmem_cache_alloc (getname+0x23) call_site=810f559d ptr=0
8 0 3280035.156997 ls 25901 ..... kmem_cache_alloc (compat_do_execve+0x43) call_site=8111c
9 1 3280035.156997 trace-cmd 25900 d.... sched_switch 25900:44:S ==> 0:140: swapper

Zoom Out

- Left click and drag left



Event Info Tool Tip

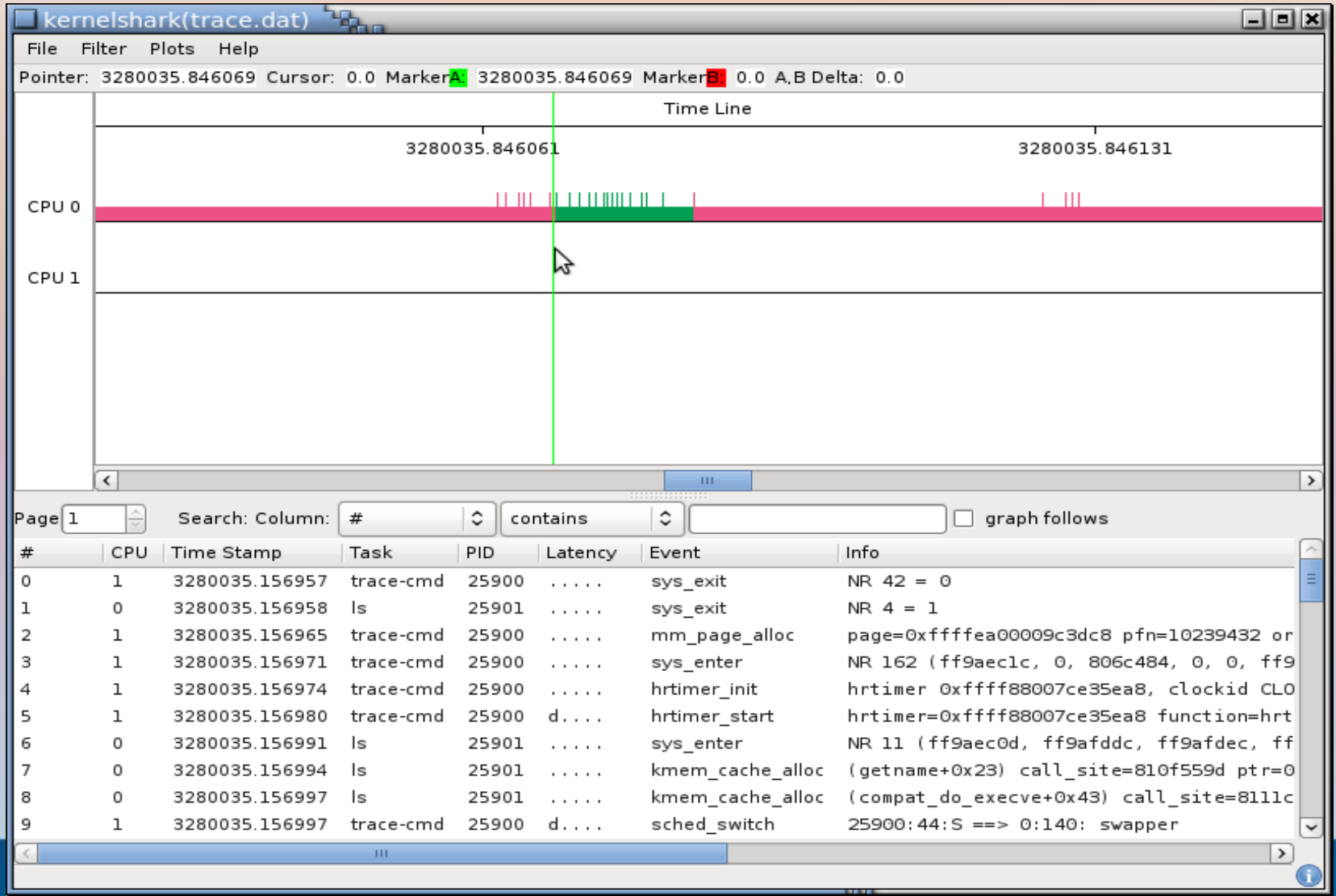


Graph Markers

- Marker A and B
- Used to calculate the deltas

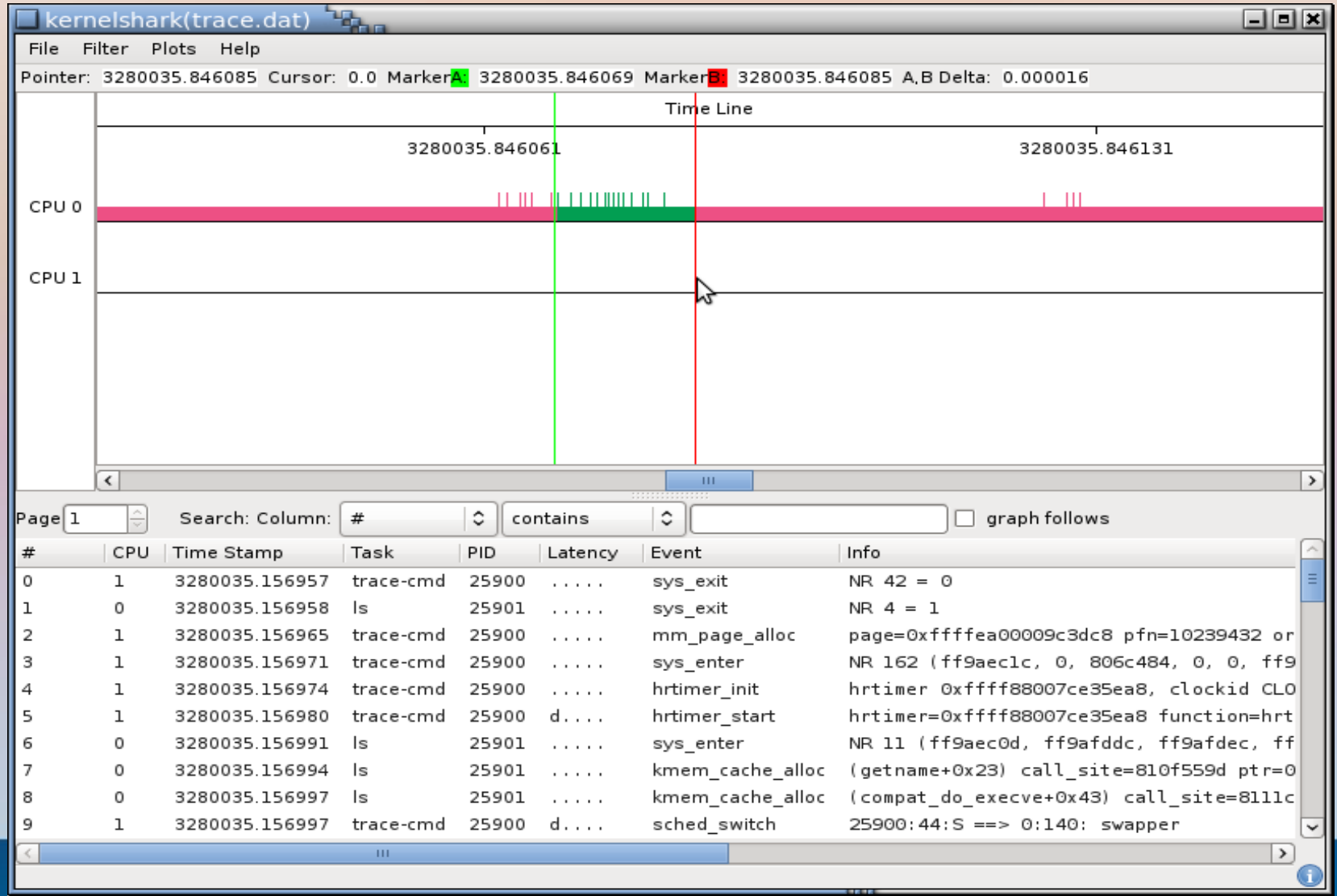
Marker A

- Left mouse click



Marker B

- Left mouse click with shift key held

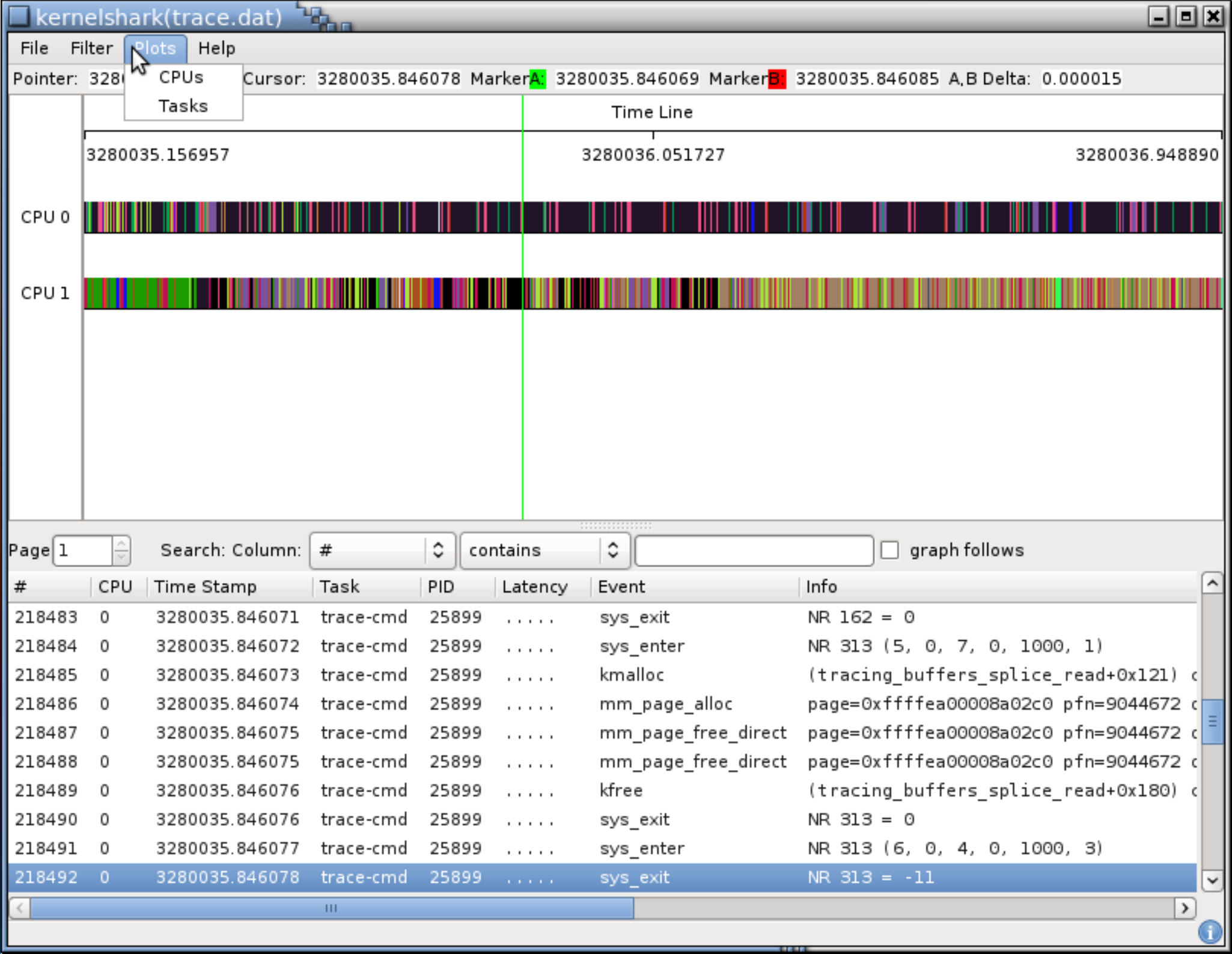


Graph Cursor

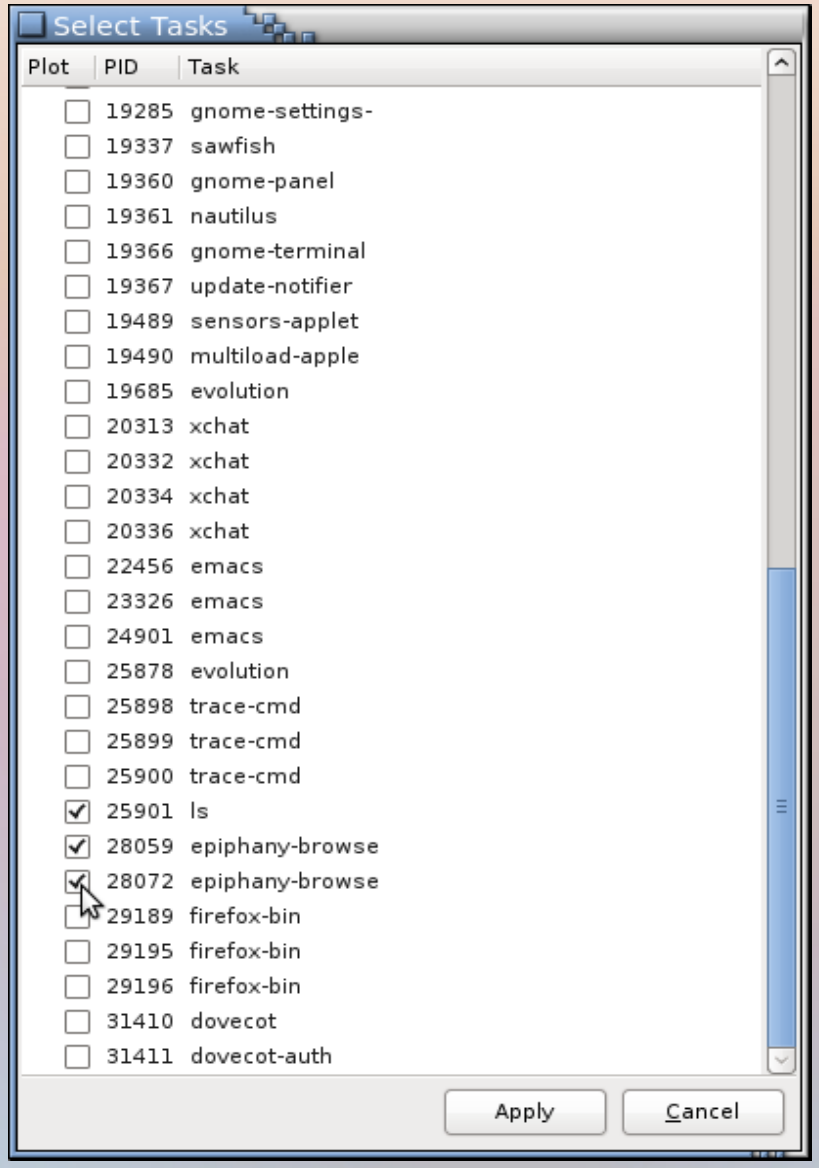
- Double click on graph
- Moves the list view to the closest event to the timestamp on where the cursor is.
- Can be used for marking location on zooming in and out

Graph Plots

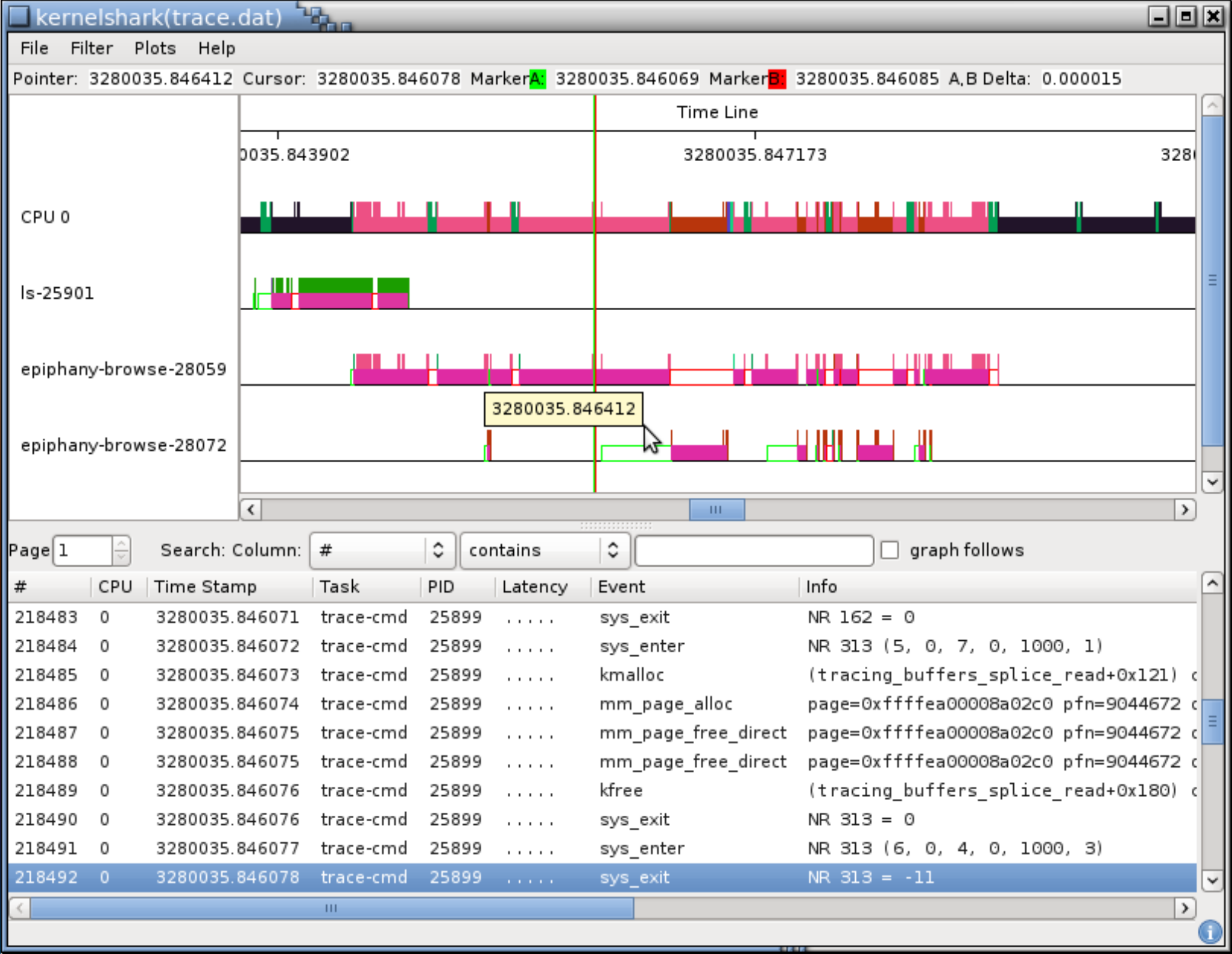
- CPU Plots
  - colors change depending on what task is running
- Task Plots
  - colors change depending on what CPU the task is on
  - shows wake up latency (hollow green box)
  - shows preempt latency (hollow red box)
  - can also be opened by menu option when mouse is over a task in the CPU plot

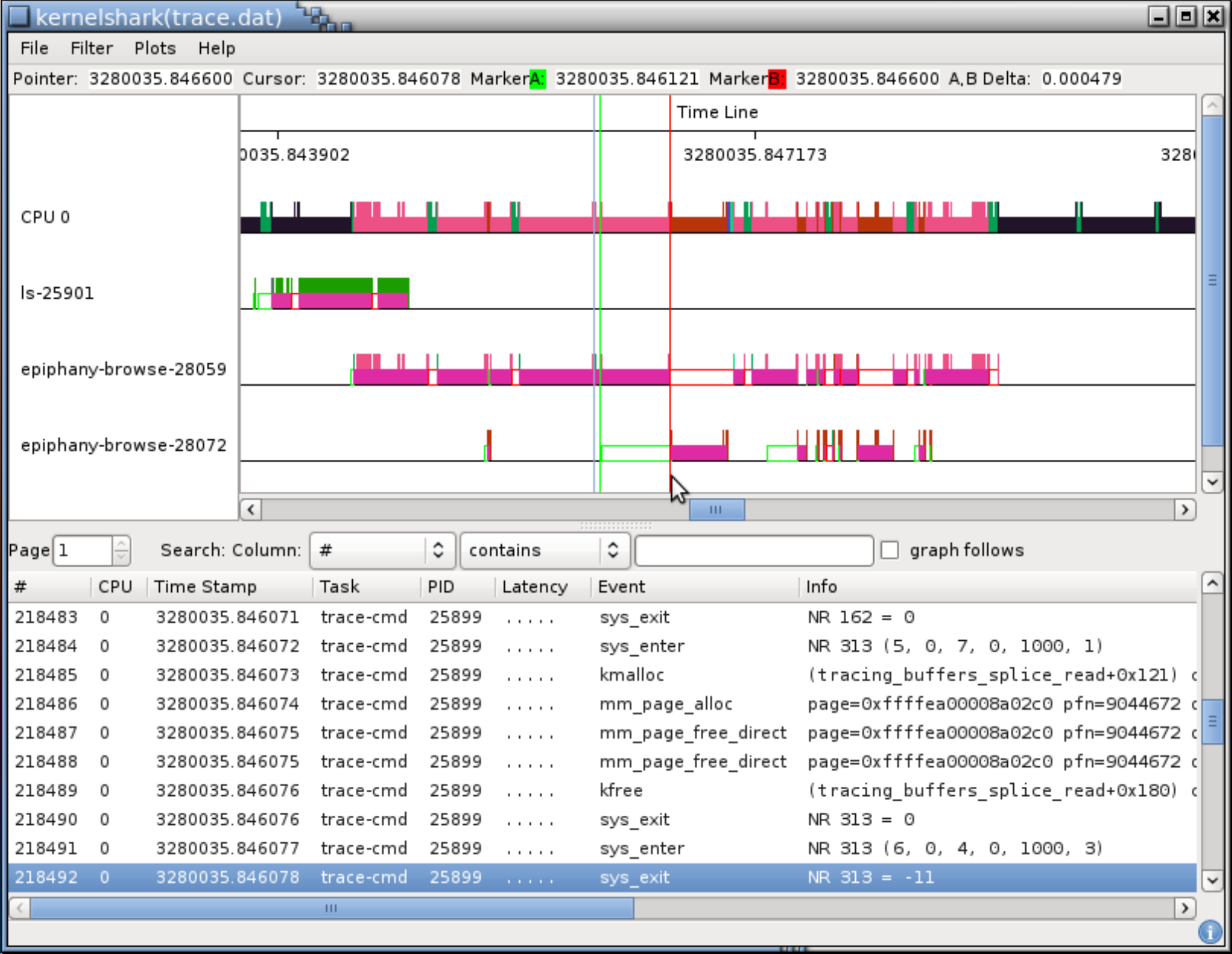

kernelshark(trace.dat)
File
Filter
Plots
Help
CPUs
Tasks
Pointer: 328
Cursor: 3280035.846078
MarkerA: 3280035.846069
MarkerB: 3280035.846085
A,B Delta: 0.000015
Time Line
3280035.156957
3280036.051727
3280036.948890
CPU 0
CPU 1
Page 1
Search: Column:
#
contains
graph follows
CPU Time Stamp Task PID Latency Event Info
218483 0 3280035.846071 trace-cmd 25899 ..... sys_exit NR 162 = 0
218484 0 3280035.846072 trace-cmd 25899 ..... sys_enter NR 313 (5, 0, 7, 0, 1000, 1)
218485 0 3280035.846073 trace-cmd 25899 ..... kmalloc (tracing_buffers_splice_read+0x121) c
218486 0 3280035.846074 trace-cmd 25899 ..... mm_page_alloc page=0xffffea00008a02c0 pfn=9044672 c
218487 0 3280035.846075 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008a02c0 pfn=9044672 c
218488 0 3280035.846075 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008a02c0 pfn=9044672 c
218489 0 3280035.846076 trace-cmd 25899 ..... kfree (tracing_buffers_splice_read+0x180) c
218490 0 3280035.846076 trace-cmd 25899 ..... sys_exit NR 313 = 0
218491 0 3280035.846077 trace-cmd 25899 ..... sys_enter NR 313 (6, 0, 4, 0, 1000, 3)
218492 0 3280035.846078 trace-cmd 25899 ..... sys_exit NR 313 = -11

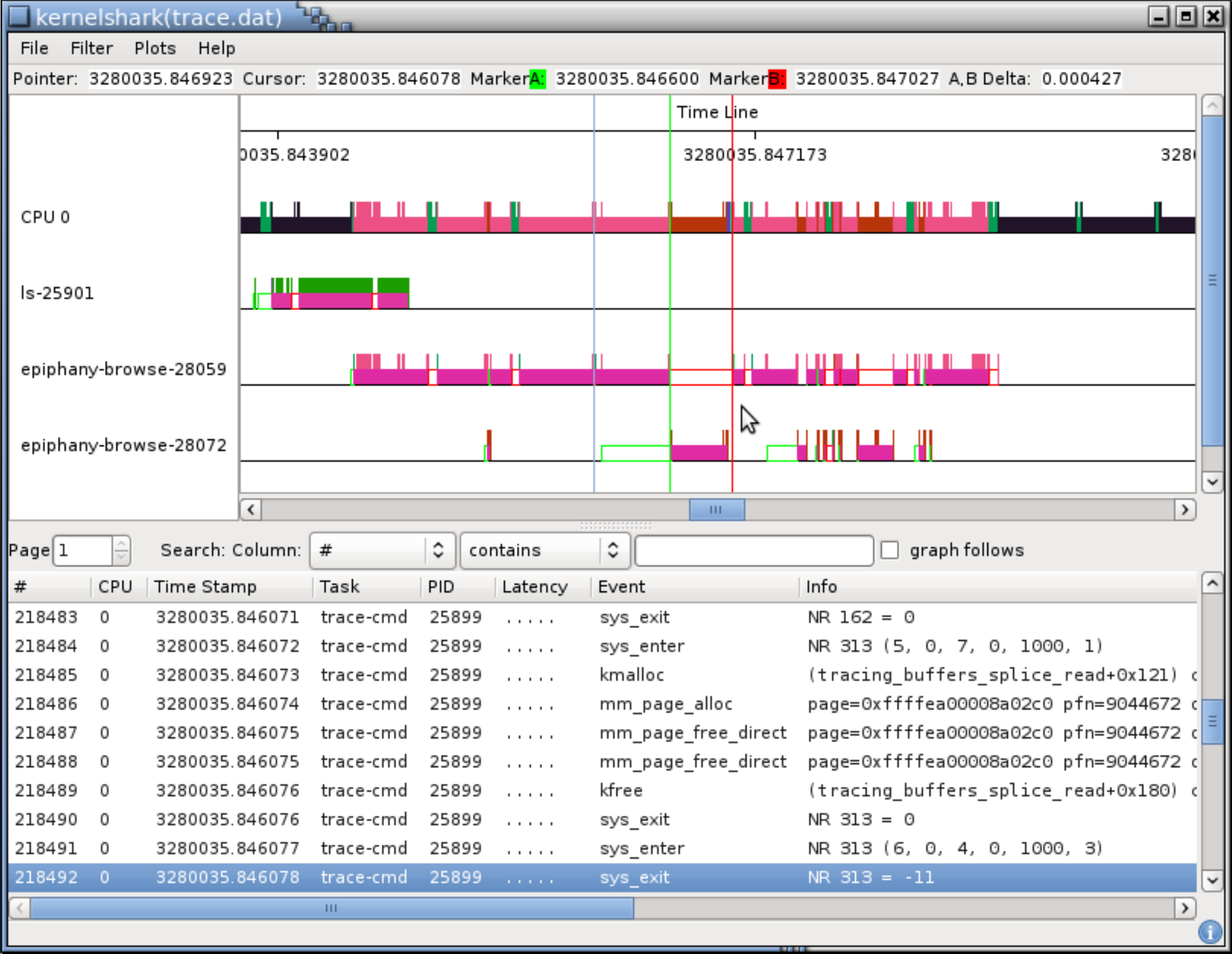
List of Tasks to plot



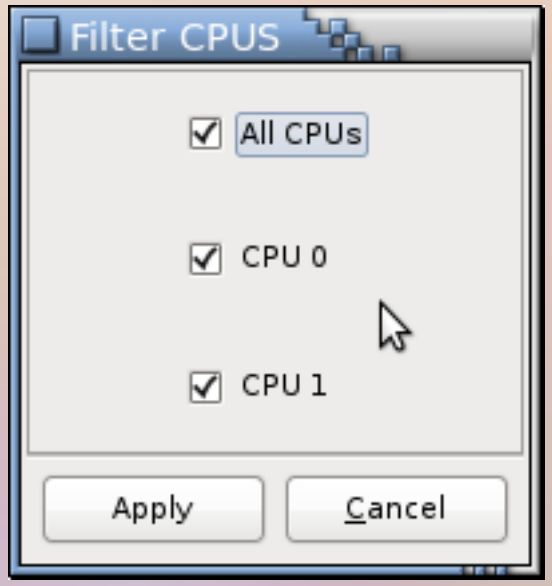
kernelshark(trace.dat)
File Filter Plots Help
Pointer: 3280036.041510 Cursor: 3280035.846078 MarkerA: 3280035.846069 MarkerB: 3280035.846085 A,B Delta: 0.000015
Time Line
3280035.156957
3280036.052923
3280036.948890
CPU 0
ls-25901
epiphany-browse-28059
epiphany-browse-28072
Page 1
Search: Column: #
contains
graph follows
CPU Time Stamp Task PID Latency Event Info
218483 0 3280035.846071 trace-cmd 25899 ..... sys_exit NR 162 = 0
218484 0 3280035.846072 trace-cmd 25899 ..... sys_enter NR 313 (5, 0, 7, 0, 1000, 1)
218485 0 3280035.846073 trace-cmd 25899 ..... kmalloc (tracing_buffers_splice_read+0x121)
218486 0 3280035.846074 trace-cmd 25899 ..... mm_page_alloc page=0xffffea00008a02c0 pfn=9044672
218487 0 3280035.846075 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008a02c0 pfn=9044672
218488 0 3280035.846075 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008a02c0 pfn=9044672
218489 0 3280035.846076 trace-cmd 25899 ..... kfree (tracing_buffers_splice_read+0x180)
218490 0 3280035.846076 trace-cmd 25899 ..... sys_exit NR 313 = 0
218491 0 3280035.846077 trace-cmd 25899 ..... sys_enter NR 313 (6, 0, 4, 0, 1000, 3)
218492 0 3280035.846078 trace-cmd 25899 ..... sys_exit NR 313 = -11

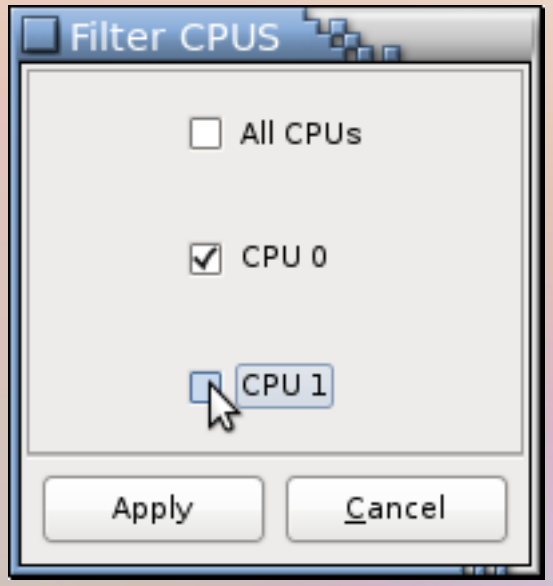

kernelshark(trace.dat)
File
Filter
Plots
Help
Pointer: 3280035.846412 Cursor: 3280035.846078 MarkerA: 3280035.846069 MarkerB: 3280035.846085 A,B Delta: 0.000015
Time Line
0035.843902
3280035.847173
CPU 0
ls-25901
epiphany-browse-28059
3280035.846412
epiphany-browse-28072
Page 1
Search: Column:
#
contains
graph follows
CPU Time Stamp Task PID Latency Event Info
218483 0 3280035.846071 trace-cmd 25899 ..... sys_exit NR 162 = 0
218484 0 3280035.846072 trace-cmd 25899 ..... sys_enter NR 313 (5, 0, 7, 0, 1000, 1)
218485 0 3280035.846073 trace-cmd 25899 ..... kmalloc (tracing_buffers_splice_read+0x121)
218486 0 3280035.846074 trace-cmd 25899 ..... mm_page_alloc page=0xffffea00008a02c0 pfn=9044672
218487 0 3280035.846075 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008a02c0 pfn=9044672
218488 0 3280035.846075 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008a02c0 pfn=9044672
218489 0 3280035.846076 trace-cmd 25899 ..... kfree (tracing_buffers_splice_read+0x180)
218490 0 3280035.846076 trace-cmd 25899 ..... sys_exit NR 313 = 0
218491 0 3280035.846077 trace-cmd 25899 ..... sys_enter NR 313 (6, 0, 4, 0, 1000, 3)
218492 0 3280035.846078 trace-cmd 25899 ..... sys_exit NR 313 = -11

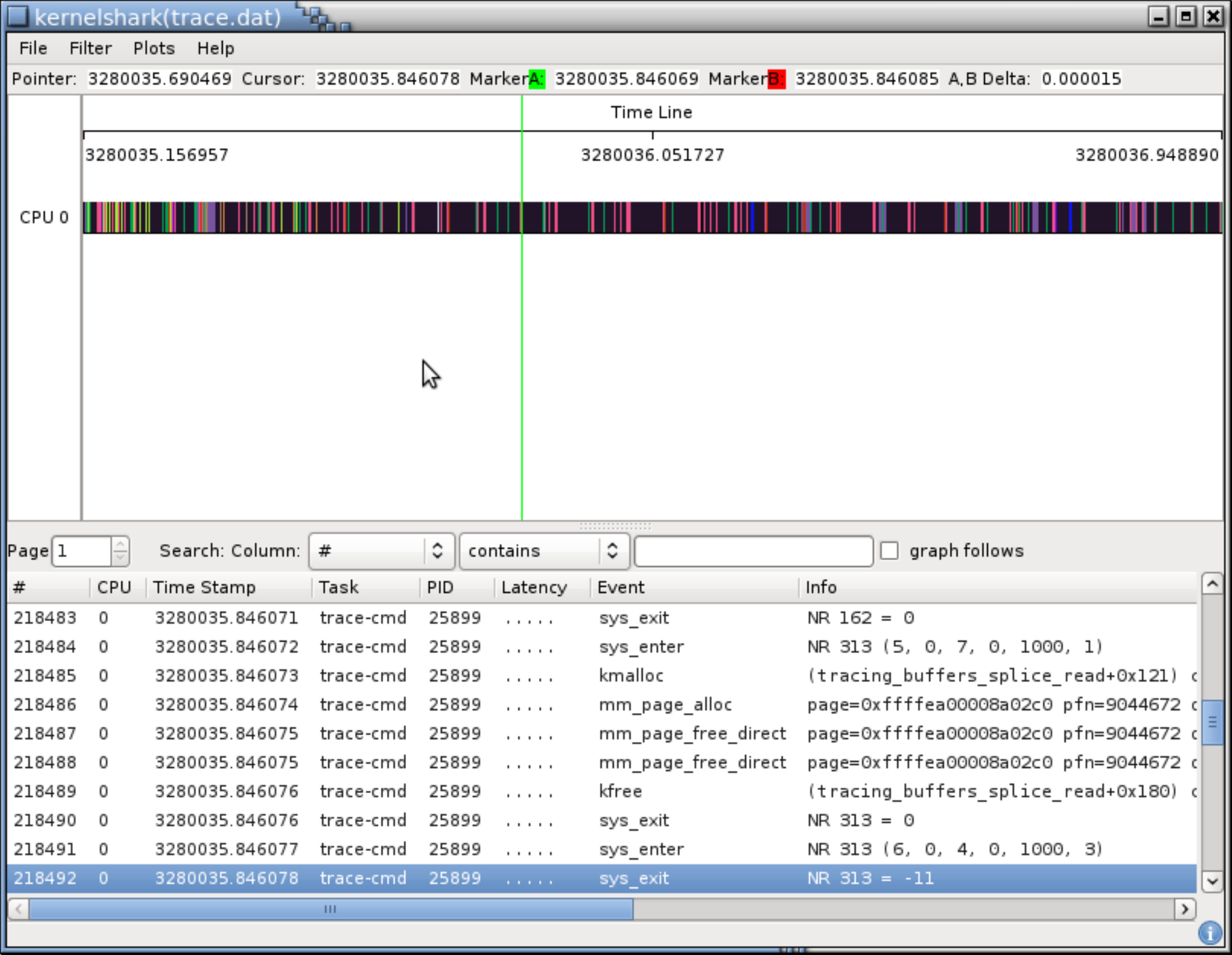

kernelshark(trace.dat)
File Filter Plots Help
Pointer: 3280035.846600 Cursor: 3280035.846078 MarkerA: 3280035.846121 MarkerB: 3280035.846600 A,B Delta: 0.000479
Time Line
0035.843902
3280035.847173
CPU 0
ls-25901
epiphany-browse-28059
epiphany-browse-28072
Page 1
Search: Column: #
contains
graph follows
CPU Time Stamp Task PID Latency Event Info
218483 0 3280035.846071 trace-cmd 25899 ..... sys_exit NR 162 = 0
218484 0 3280035.846072 trace-cmd 25899 ..... sys_enter NR 313 (5, 0, 7, 0, 1000, 1)
218485 0 3280035.846073 trace-cmd 25899 ..... kmalloc (tracing_buffers_splice_read+0x121)
218486 0 3280035.846074 trace-cmd 25899 ..... mm_page_alloc page=0xffffea00008a02c0 pfn=9044672
218487 0 3280035.846075 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008a02c0 pfn=9044672
218488 0 3280035.846075 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008a02c0 pfn=9044672
218489 0 3280035.846076 trace-cmd 25899 ..... kfree (tracing_buffers_splice_read+0x180)
218490 0 3280035.846076 trace-cmd 25899 ..... sys_exit NR 313 = 0
218491 0 3280035.846077 trace-cmd 25899 ..... sys_enter NR 313 (6, 0, 4, 0, 1000, 3)
218492 0 3280035.846078 trace-cmd 25899 ..... sys_exit NR 313 = -11


kernelshark(trace.dat)
File Filter Plots Help
Pointer: 3280035.846923 Cursor: 3280035.846078 MarkerA: 3280035.846600 MarkerB: 3280035.847027 A,B Delta: 0.000427
Time Line
0035.843902
3280035.847173
CPU 0
ls-25901
epiphany-browse-28059
epiphany-browse-28072
Page 1 Search: Column: # contains graph follows
CPU Time Stamp Task PID Latency Event Info
218483 0 3280035.846071 trace-cmd 25899 ..... sys_exit NR 162 = 0
218484 0 3280035.846072 trace-cmd 25899 ..... sys_enter NR 313 (5, 0, 7, 0, 1000, 1)
218485 0 3280035.846073 trace-cmd 25899 ..... kmalloc (tracing_buffers_splice_read+0x121)
218486 0 3280035.846074 trace-cmd 25899 ..... mm_page_alloc page=0xffffea00008a02c0 pfn=9044672
218487 0 3280035.846075 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008a02c0 pfn=9044672
218488 0 3280035.846075 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008a02c0 pfn=9044672
218489 0 3280035.846076 trace-cmd 25899 ..... kfree (tracing_buffers_splice_read+0x180)
218490 0 3280035.846076 trace-cmd 25899 ..... sys_exit NR 313 = 0
218491 0 3280035.846077 trace-cmd 25899 ..... sys_enter NR 313 (6, 0, 4, 0, 1000, 3)
218492 0 3280035.846078 trace-cmd 25899 ..... sys_exit NR 313 = -11

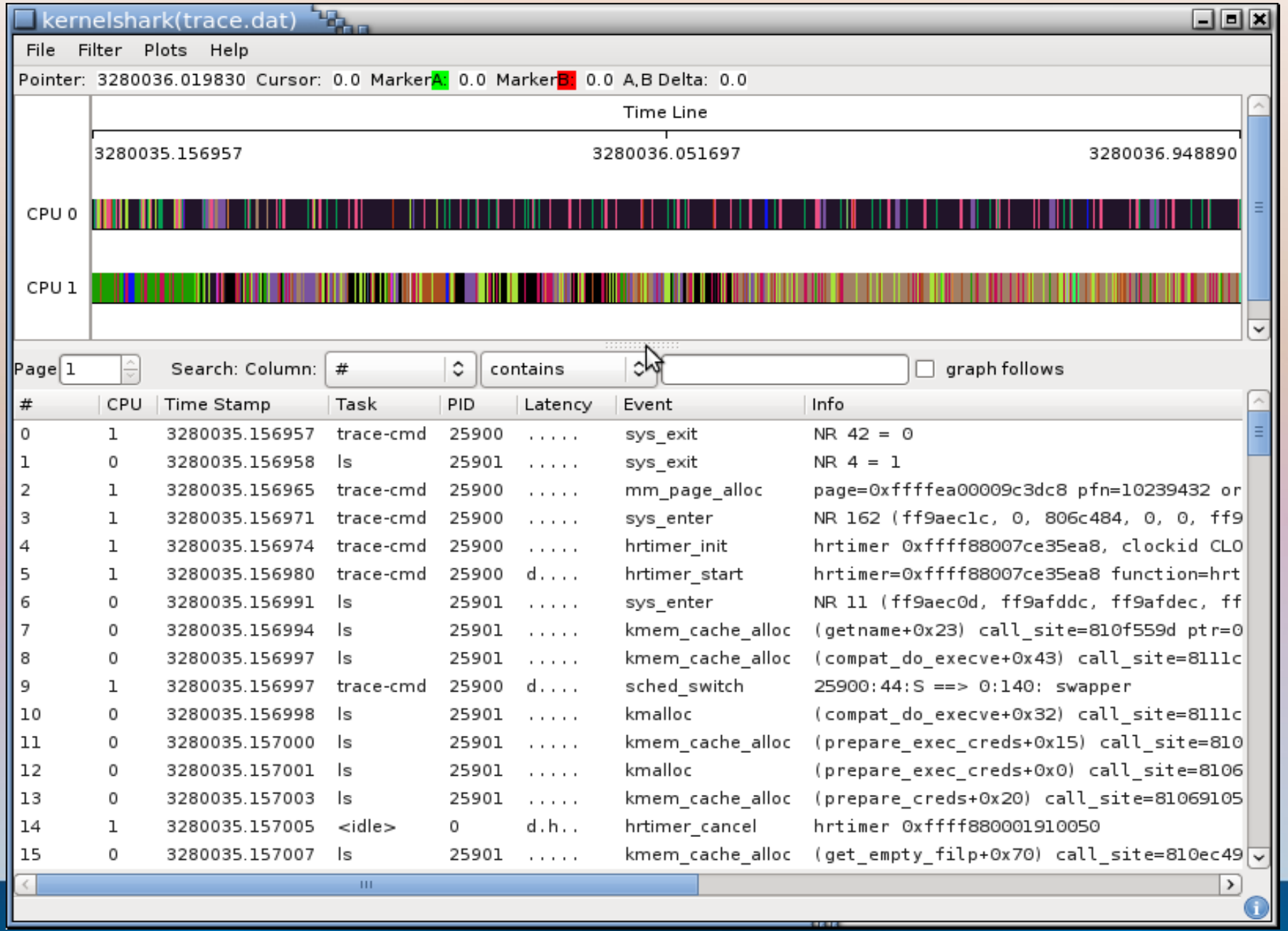
CPU Plots


Filter CPUS
All CPUs
CPU 0
CPU 1
Apply
Cancel


Filter CPUS
All CPUs
CPU 0
CPU 1
Apply
Cancel


kernelshark(trace.dat)
File
Filter
Plots
Help
Pointer: 3280035.690469 Cursor: 3280035.846078 MarkerA: 3280035.846069 MarkerB: 3280035.846085 A,B Delta: 0.000015
Time Line
3280035.156957
3280036.051727
3280036.948890
CPU 0
Page 1
Search: Column: #
contains
graph follows
CPU Time Stamp Task PID Latency Event Info
218483 0 3280035.846071 trace-cmd 25899 ..... sys_exit NR 162 = 0
218484 0 3280035.846072 trace-cmd 25899 ..... sys_enter NR 313 (5, 0, 7, 0, 1000, 1)
218485 0 3280035.846073 trace-cmd 25899 ..... kmalloc (tracing_buffers_splice_read+0x121) c
218486 0 3280035.846074 trace-cmd 25899 ..... mm_page_alloc page=0xffffea00008a02c0 pfn=9044672 c
218487 0 3280035.846075 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008a02c0 pfn=9044672 c
218488 0 3280035.846075 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008a02c0 pfn=9044672 c
218489 0 3280035.846076 trace-cmd 25899 ..... kfree (tracing_buffers_splice_read+0x180) c
218490 0 3280035.846076 trace-cmd 25899 ..... sys_exit NR 313 = 0
218491 0 3280035.846077 trace-cmd 25899 ..... sys_enter NR 313 (6, 0, 4, 0, 1000, 3)
218492 0 3280035.846078 trace-cmd 25899 ..... sys_exit NR 313 = -11

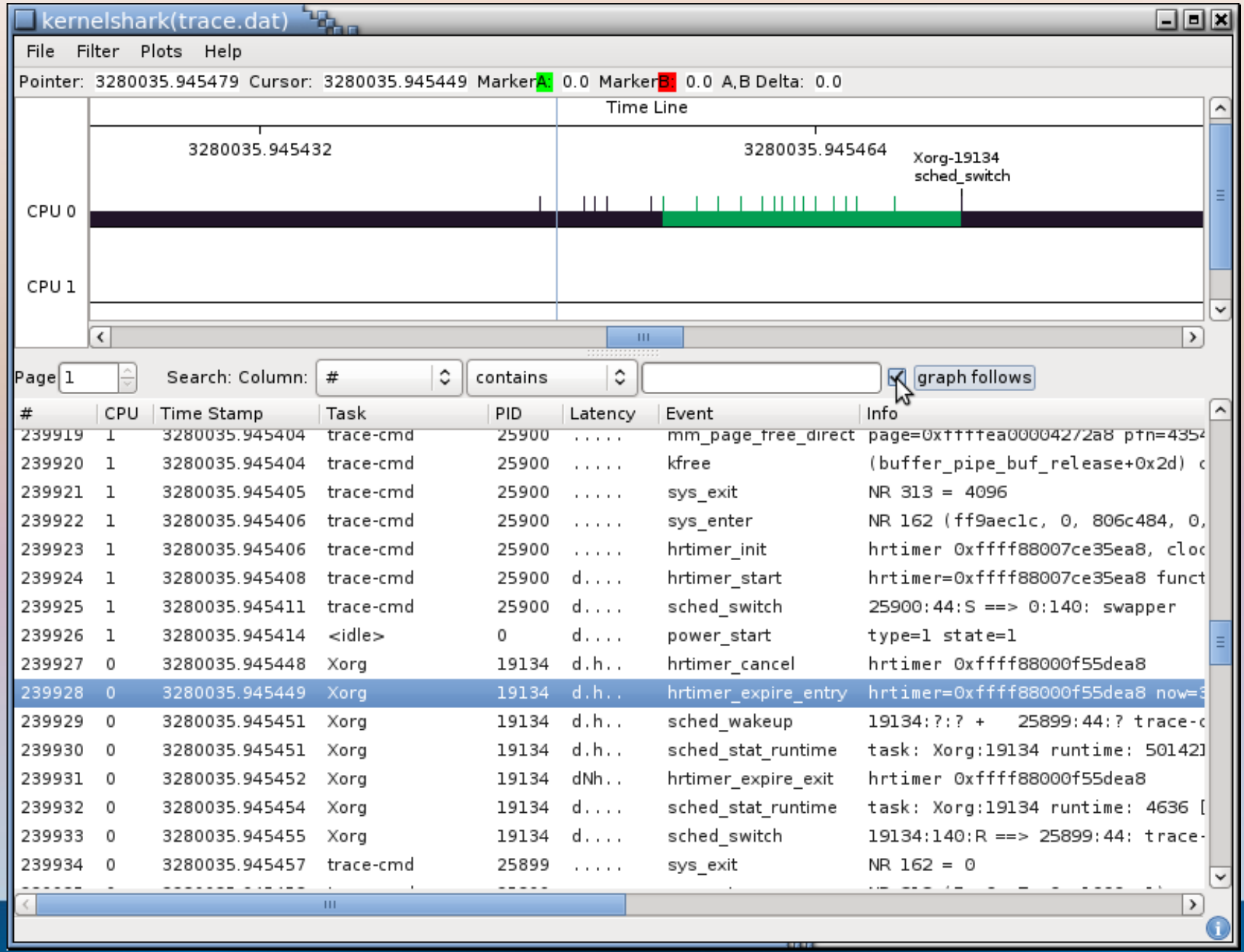
List view


kernelshark(trace.dat)
File Filter Plots Help
Pointer: 3280036.019830 Cursor: 0.0 MarkerA: 0.0 MarkerB: 0.0 A,B Delta: 0.0
Time Line
3280035.156957
3280036.051697
3280036.948890
CPU 0
CPU 1
Page 1
Search: Column: #
contains
graph follows
CPU Time Stamp Task PID Latency Event Info
0 1 3280035.156957 trace-cmd 25900 ..... sys_exit NR 42 = 0
1 0 3280035.156958 ls 25901 ..... sys_exit NR 4 = 1
2 1 3280035.156965 trace-cmd 25900 ..... mm_page_alloc page=0xffffea00009c3dc8 pfn=10239432 or
3 1 3280035.156971 trace-cmd 25900 ..... sys_enter NR 162 (ff9aec1c, 0, 806c484, 0, 0, ff9
4 1 3280035.156974 trace-cmd 25900 ..... hrtimer_init hrtimer 0xffff88007ce35ea8, clockid CLO
5 1 3280035.156980 trace-cmd 25900 d.... hrtimer_start hrtimer=0xffff88007ce35ea8 function=hrt
6 0 3280035.156991 ls 25901 ..... sys_enter NR 11 (ff9aec0d, ff9afddc, ff9afdec, ff
7 0 3280035.156994 ls 25901 ..... kmem_cache_alloc (getname+0x23) call_site=810f559d ptr=0
8 0 3280035.156997 ls 25901 ..... kmem_cache_alloc (compat_do_execve+0x43) call_site=8111c
9 1 3280035.156997 trace-cmd 25900 d.... sched_switch 25900:44:S ==> 0:140: swapper
10 0 3280035.156998 ls 25901 ..... kmalloc (compat_do_execve+0x32) call_site=8111c
11 0 3280035.157000 ls 25901 ..... kmem_cache_alloc (prepare_exec_creds+0x15) call_site=810
12 0 3280035.157001 ls 25901 ..... kmalloc (prepare_exec_creds+0x0) call_site=8106
13 0 3280035.157003 ls 25901 ..... kmem_cache_alloc (prepare_creds+0x20) call_site=81069105
14 1 3280035.157005 <idle> 0 d.h.. hrtimer_cancel hrtimer 0xffff880001910050
15 0 3280035.157007 ls 25901 ..... kmem_cache_alloc (get_empty_filp+0x70) call_site=810ec49

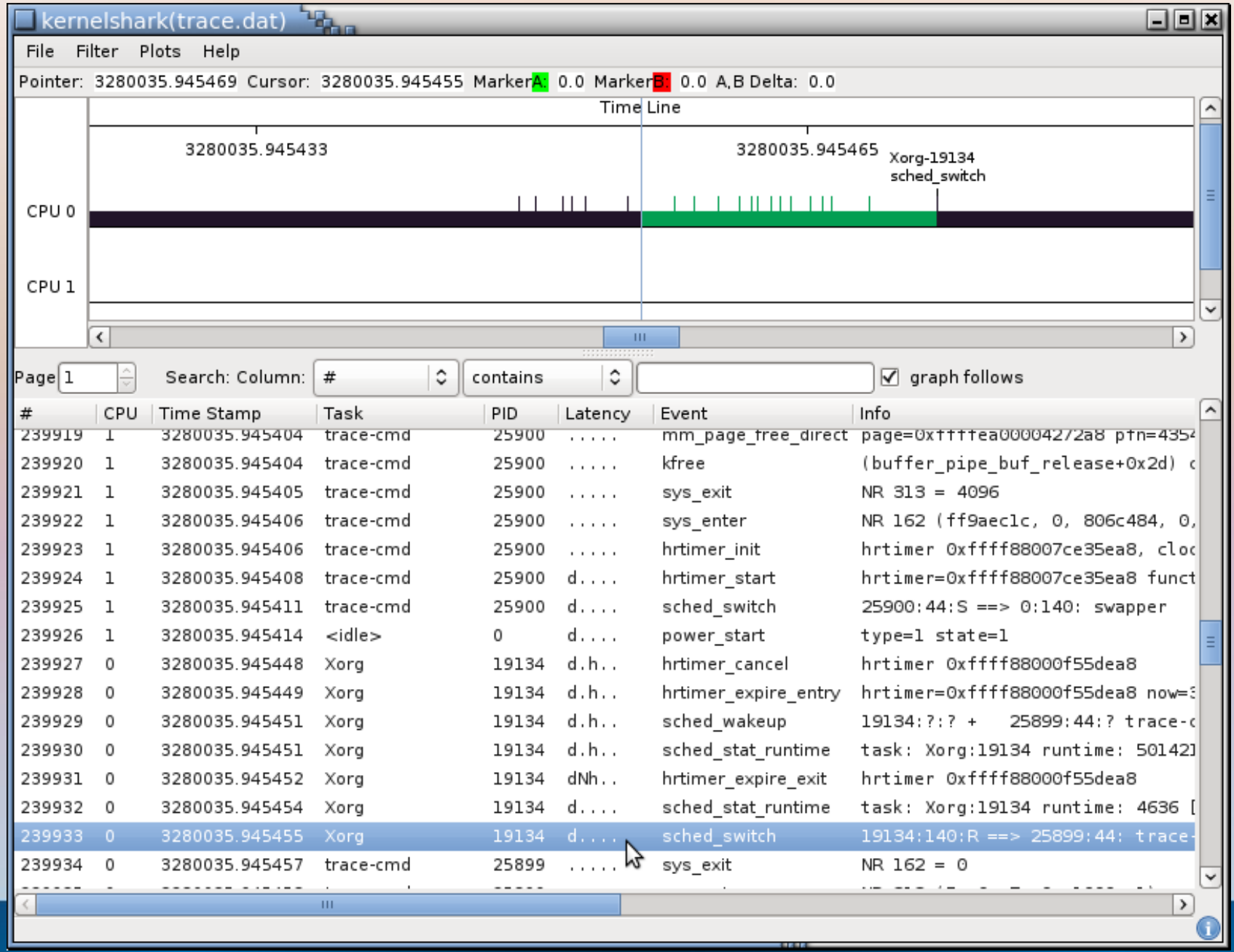
Search the List

- Search by column
  - Contains
  - Full match
  - Does not have

Graph follows toggle

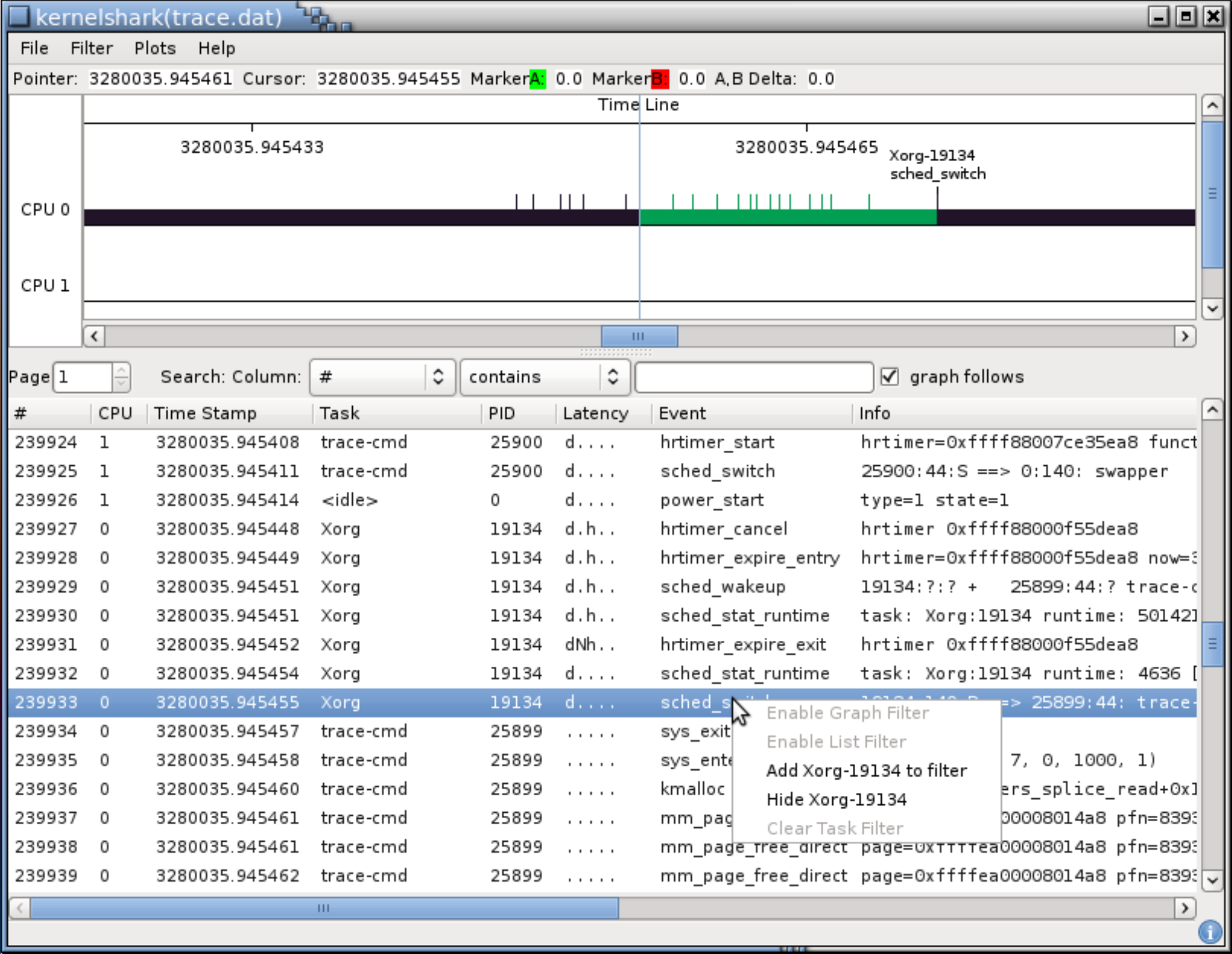


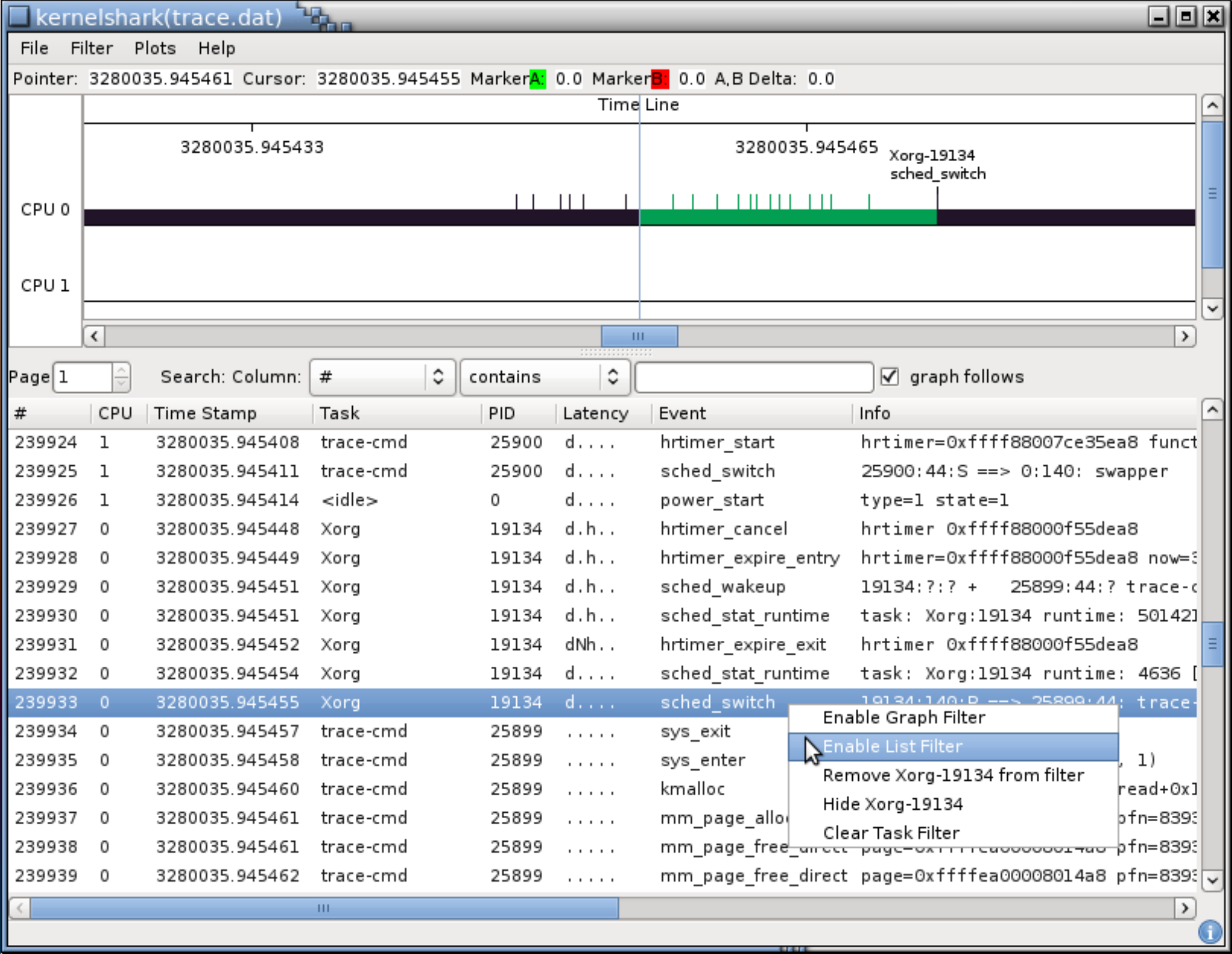
Graph follows toggle



Filtering

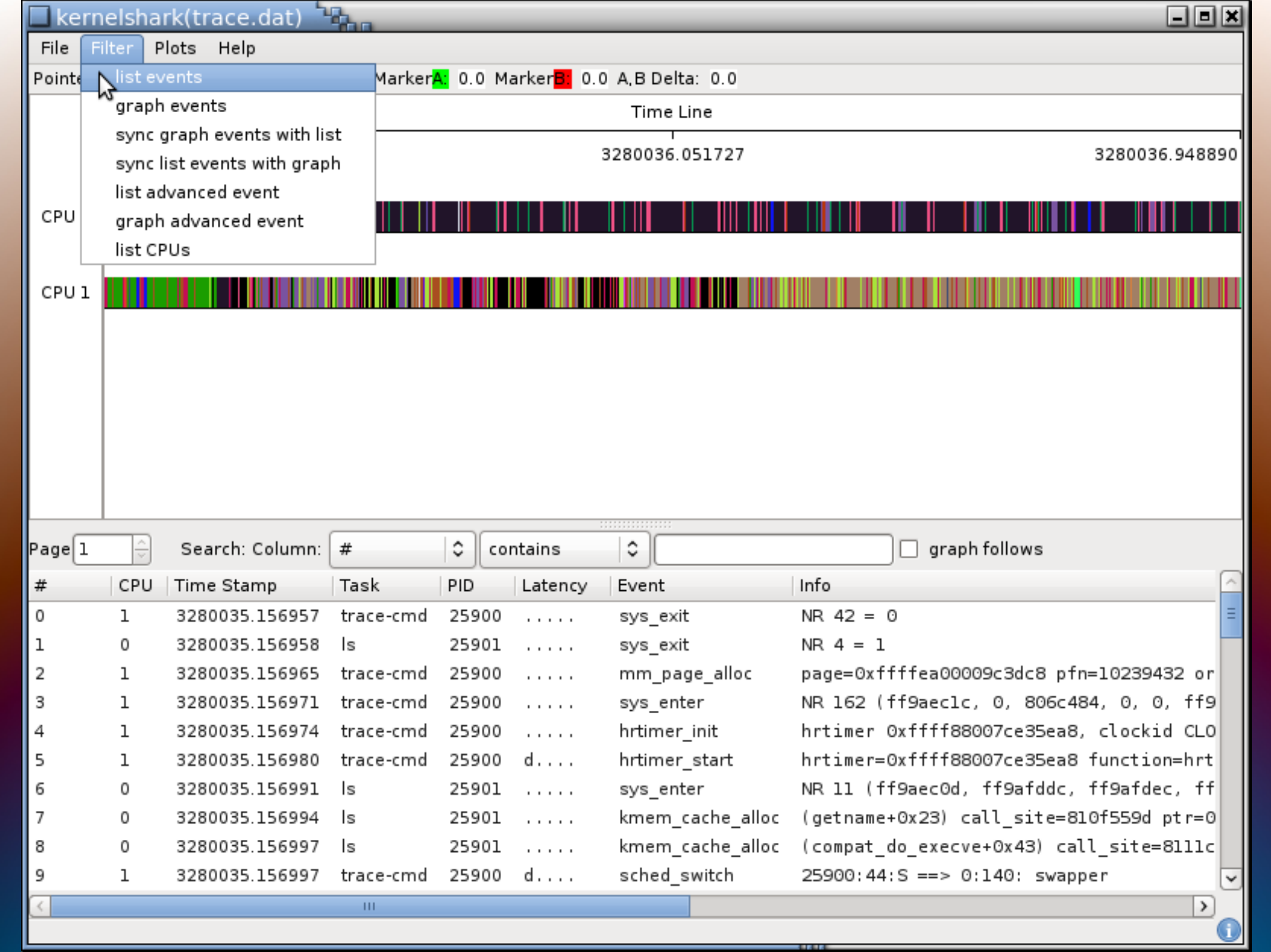
- Filter out tasks
- Filter in tasks
- Filter events
- Filter events based on content


kernelshark(trace.dat)
File Filter Plots Help
Pointer: 3280035.945461 Cursor: 3280035.945455 MarkerA: 0.0 MarkerB: 0.0 A,B Delta: 0.0
Time Line
3280035.945433
3280035.945465
Xorg-19134
sched_switch
CPU 0
CPU 1
Page 1 Search: Column: # contains graph follows
CPU Time Stamp Task PID Latency Event Info
239924 1 3280035.945408 trace-cmd 25900 d.... hrtimer_start hrtimer=0xffff88007ce35ea8 funct
239925 1 3280035.945411 trace-cmd 25900 d.... sched_switch 25900:44:S ==> 0:140: swapper
239926 1 3280035.945414 <idle> 0 d.... power_start type=1 state=1
239927 0 3280035.945448 Xorg 19134 d.h.. hrtimer_cancel hrtimer 0xffff88000f55dea8
239928 0 3280035.945449 Xorg 19134 d.h.. hrtimer_expire_entry hrtimer=0xffff88000f55dea8 now=3
239929 0 3280035.945451 Xorg 19134 d.h.. sched_wakeup 19134:?:? + 25899:44:? trace-c
239930 0 3280035.945451 Xorg 19134 d.h.. sched_stat_runtime task: Xorg:19134 runtime: 501421
239931 0 3280035.945452 Xorg 19134 dNh.. hrtimer_expire_exit hrtimer 0xffff88000f55dea8
239932 0 3280035.945454 Xorg 19134 d.... sched_stat_runtime task: Xorg:19134 runtime: 4636 [
239933 0 3280035.945455 Xorg 19134 d.... sched_s => 25899:44: trace-
239934 0 3280035.945457 trace-cmd 25899 ..... sys_exit
239935 0 3280035.945458 trace-cmd 25899 ..... sys_ente 7, 0, 1000, 1)
239936 0 3280035.945460 trace-cmd 25899 ..... kmalloc ers_splice_read+0x1
239937 0 3280035.945461 trace-cmd 25899 ..... mm_pag 00008014a8 pfn=8393
239938 0 3280035.945461 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008014a8 pfn=8393
239939 0 3280035.945462 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008014a8 pfn=8393
Enable Graph Filter
Enable List Filter
Add Xorg-19134 to filter
Hide Xorg-19134
Clear Task Filter

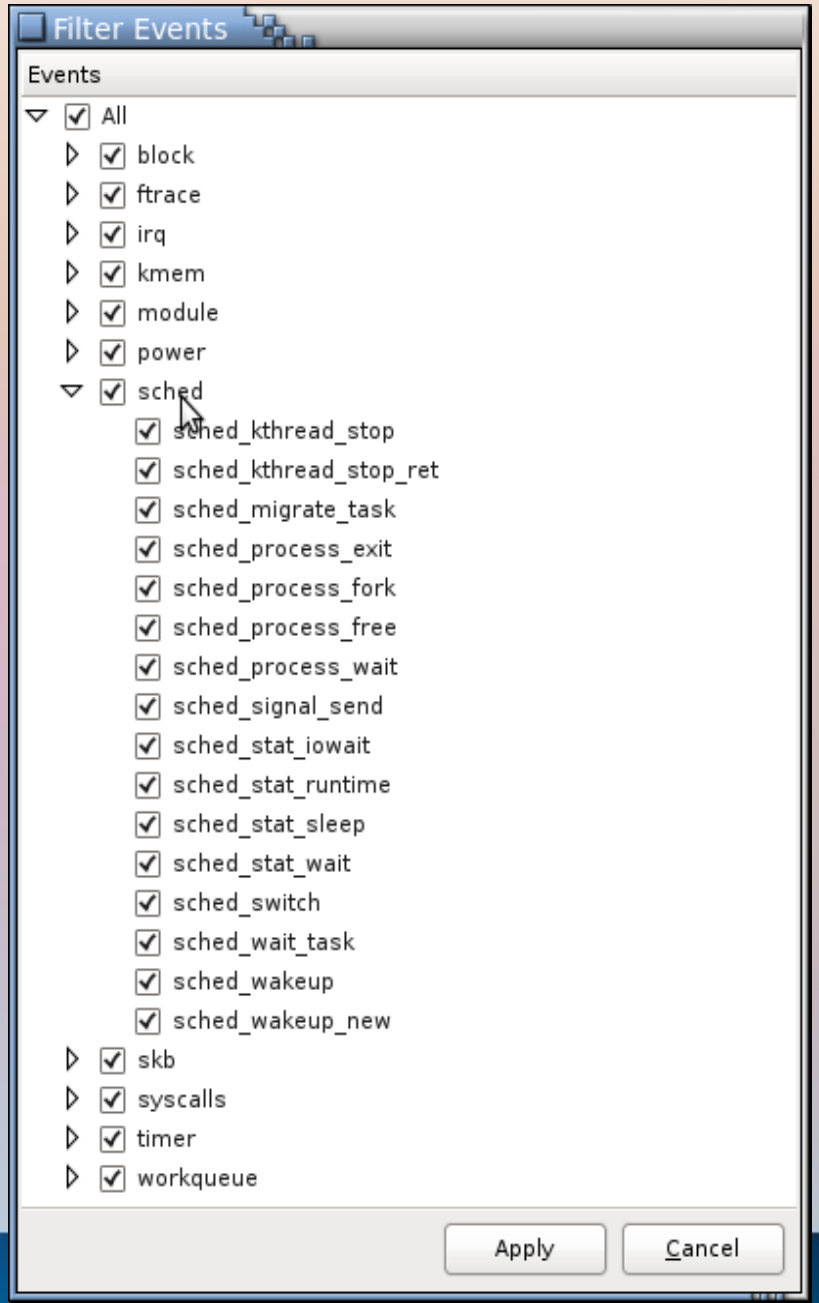

kernelshark(trace.dat)
File
Filter
Plots
Help
Pointer: 3280035.945461 Cursor: 3280035.945455 MarkerA: 0.0 MarkerB: 0.0 A,B Delta: 0.0
Time Line
3280035.945433
3280035.945465
Xorg-19134
sched_switch
CPU 0
CPU 1
Page 1
Search: Column: #
contains
graph follows
CPU Time Stamp Task PID Latency Event Info
239924 1 3280035.945408 trace-cmd 25900 d.... hrtimer_start hrtimer=0xffff88007ce35ea8 funct
239925 1 3280035.945411 trace-cmd 25900 d.... sched_switch 25900:44:S ==> 0:140: swapper
239926 1 3280035.945414 <idle> 0 d.... power_start type=1 state=1
239927 0 3280035.945448 Xorg 19134 d.h.. hrtimer_cancel hrtimer 0xffff88000f55dea8
239928 0 3280035.945449 Xorg 19134 d.h.. hrtimer_expire_entry hrtimer=0xffff88000f55dea8 now=3
239929 0 3280035.945451 Xorg 19134 d.h.. sched_wakeup 19134:?:? + 25899:44:? trace-c
239930 0 3280035.945451 Xorg 19134 d.h.. sched_stat_runtime task: Xorg:19134 runtime: 501421
239931 0 3280035.945452 Xorg 19134 dNh.. hrtimer_expire_exit hrtimer 0xffff88000f55dea8
239932 0 3280035.945454 Xorg 19134 d.... sched_stat_runtime task: Xorg:19134 runtime: 4636 [
239933 0 3280035.945455 Xorg 19134 d.... sched_switch
239934 0 3280035.945457 trace-cmd 25899 ..... sys_exit
239935 0 3280035.945458 trace-cmd 25899 ..... sys_enter
239936 0 3280035.945460 trace-cmd 25899 ..... kmalloc
239937 0 3280035.945461 trace-cmd 25899 .....
239938 0 3280035.945461 trace-cmd 25899 .....
239939 0 3280035.945462 trace-cmd 25899 ..... mm_page_free_direct page=0xffffea00008014a8 pfn=8393
Enable Graph Filter
Enable List Filter
Remove Xorg-19134 from filter
Hide Xorg-19134
Clear Task Filter

Scheduling events

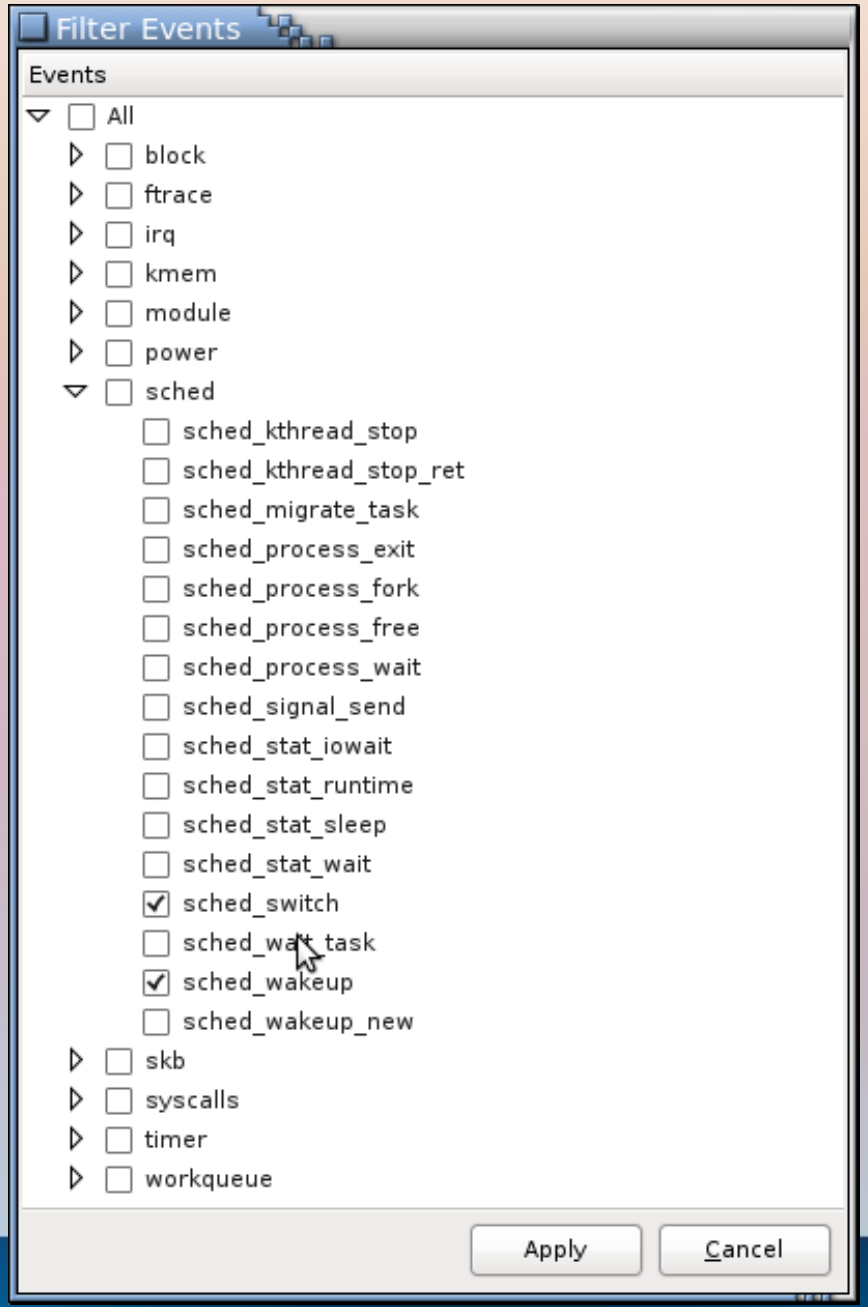
- sched_switch
- sched_wakeup
- sched_wakeup_new
- If a task in either side is to be displayed, then the event will be displayed

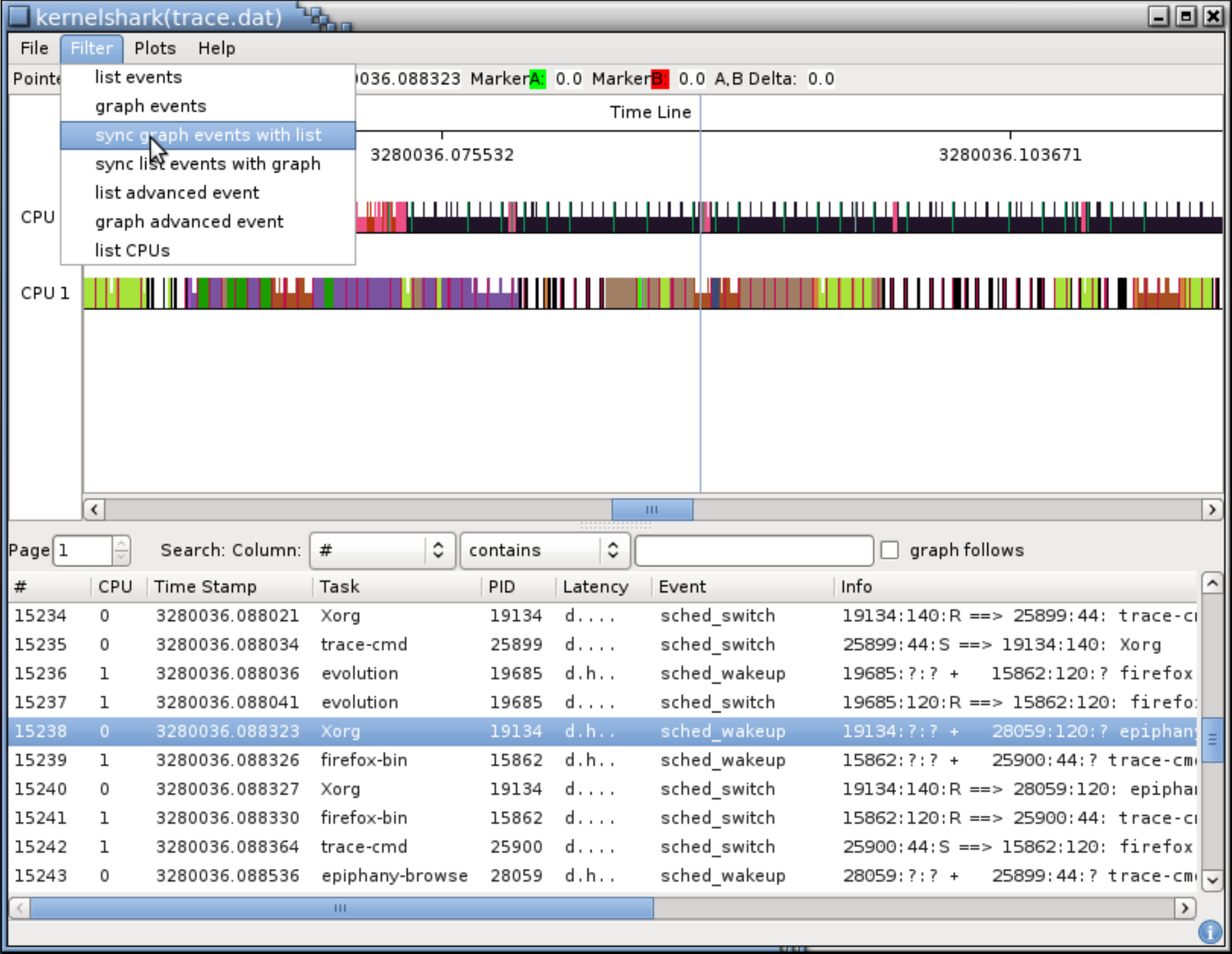

kernelshark(trace.dat)
File
Filter
Plots
Help
list events
graph events
sync graph events with list
sync list events with graph
list advanced event
graph advanced event
list CPUs
MarkerA: 0.0 MarkerB: 0.0 A,B Delta: 0.0
Time Line
3280036.051727
3280036.948890
CPU
CPU 1
Page 1
Search: Column: #
contains
graph follows
CPU Time Stamp Task PID Latency Event Info
0 1 3280035.156957 trace-cmd 25900 ..... sys_exit NR 42 = 0
1 0 3280035.156958 ls 25901 ..... sys_exit NR 4 = 1
2 1 3280035.156965 trace-cmd 25900 ..... mm_page_alloc page=0xffffea00009c3dc8 pfn=10239432 or
3 1 3280035.156971 trace-cmd 25900 ..... sys_enter NR 162 (ff9aec1c, 0, 806c484, 0, 0, ff9
4 1 3280035.156974 trace-cmd 25900 ..... hrtimer_init hrtimer 0xffff88007ce35ea8, clockid CLO
5 1 3280035.156980 trace-cmd 25900 d.... hrtimer_start hrtimer=0xffff88007ce35ea8 function=hrt
6 0 3280035.156991 ls 25901 ..... sys_enter NR 11 (ff9aec0d, ff9afddc, ff9afdec, ff
7 0 3280035.156994 ls 25901 ..... kmem_cache_alloc (getname+0x23) call_site=810f559d ptr=0
8 0 3280035.156997 ls 25901 ..... kmem_cache_alloc (compat_do_execve+0x43) call_site=8111c
9 1 3280035.156997 trace-cmd 25900 d.... sched_switch 25900:44:S ==> 0:140: swapper

Event Filters

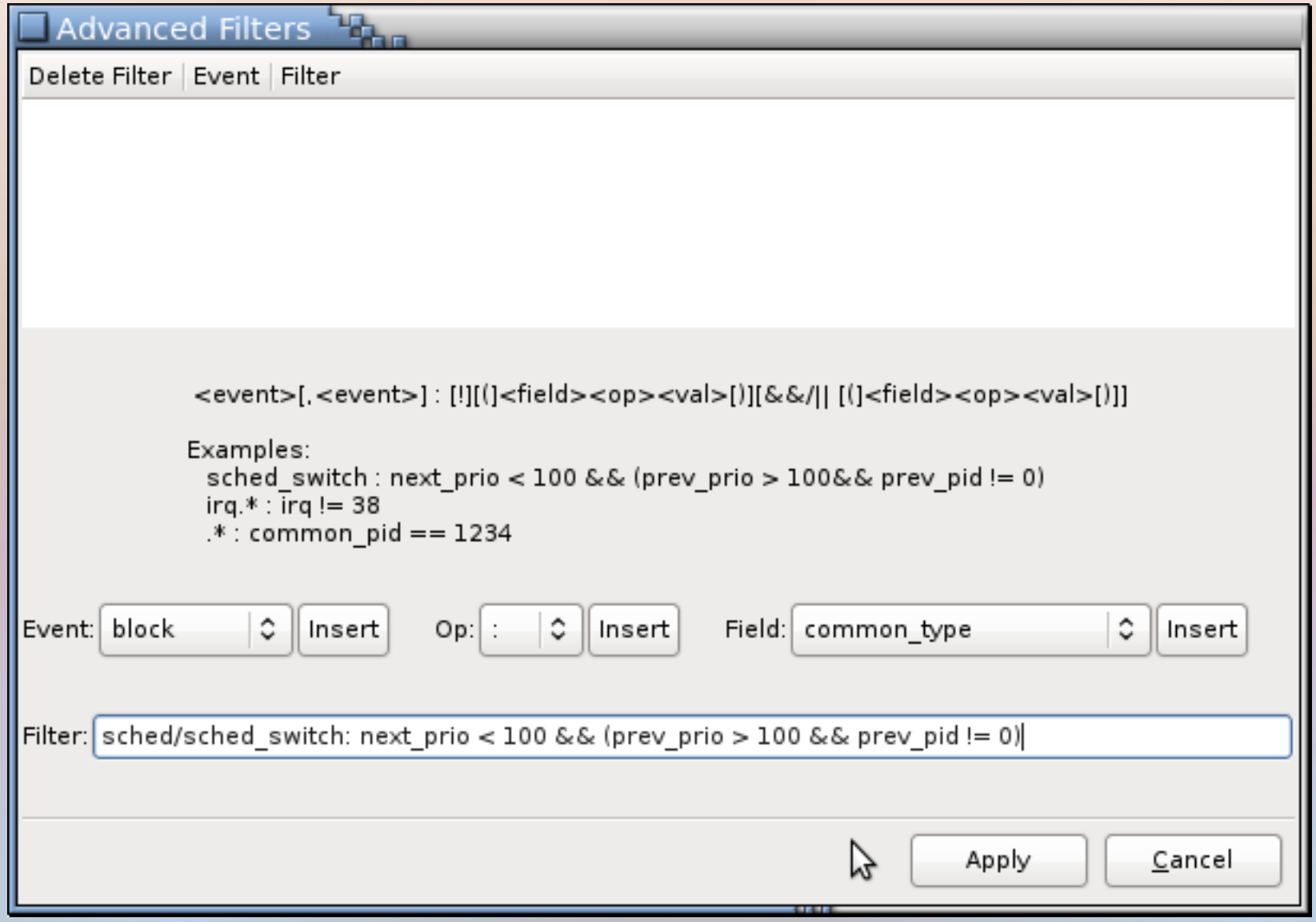


Event Filters




kernelshark(trace.dat)
File
Filter
Plots
Help
list events
graph events
sync graph events with list
sync list events with graph
list advanced event
graph advanced event
list CPUs
Pointe
036.088323 MarkerA: 0.0 MarkerB: 0.0 A,B Delta: 0.0
Time Line
3280036.075532
3280036.103671
CPU
CPU 1
Page 1
Search: Column:
#
contains
graph follows
CPU Time Stamp Task PID Latency Event Info
15234 0 3280036.088021 Xorg 19134 d.... sched_switch 19134:140:R ==> 25899:44: trace-cr
15235 0 3280036.088034 trace-cmd 25899 d.... sched_switch 25899:44:S ==> 19134:140: Xorg
15236 1 3280036.088036 evolution 19685 d.h.. sched_wakeup 19685:?:? + 15862:120:? firefox
15237 1 3280036.088041 evolution 19685 d.... sched_switch 19685:120:R ==> 15862:120: firefo
15238 0 3280036.088323 Xorg 19134 d.h.. sched_wakeup 19134:?:? + 28059:120:? epiphan
15239 1 3280036.088326 firefox-bin 15862 d.h.. sched_wakeup 15862:?:? + 25900:44:? trace-cm
15240 0 3280036.088327 Xorg 19134 d.... sched_switch 19134:140:R ==> 28059:120: epipha
15241 1 3280036.088330 firefox-bin 15862 d.... sched_switch 15862:120:R ==> 25900:44: trace-cr
15242 1 3280036.088364 trace-cmd 25900 d.... sched_switch 25900:44:S ==> 15862:120: firefox
15243 0 3280036.088536 epiphany-browse 28059 d.h.. sched_wakeup 28059:?:? + 25899:44:? trace-cm

Advanced Event Filtering



Advanced Filtering Language

```
FILTER := EVENTS | EVENTS ':' EXPRESSION
EVENTS := EVENTS ',' EVENTS | SYSTEM '/' EVENT | SYSTEM | EVENT
SYSTEM := any system name
EVENT  := any event name
EXPRESSION := EXPRESSION BOOL EXPRESSION | '(' EXPRESSION ')' | OPERATION
BOOL   := '&&' | '||'
OPERATION := '!' EXPRESSION | LVALUE CMP RVALUE | LVALUE STRCMP STRVALUE
CMP    := '>' | '<' | '==' | '>=' | '<=' | '!='
STRCMP := '==' | '!=' | '=~' | '!~'
RVALUE := integer | FIELD
STRVALUE := string (double quoted value) | FIELD
LVALUE := FIELD | EXPR
EXPR   := FIELD OP RVALUE | '(' EXPR ')' | EXPR OP EXPR
FIELD  := a field name of an event
OP     := '+' | '-' | '*' | '/' | '<<' | '>>' | '&' | '!'
```

Fields not in Events

- Field not in an event evaluates the local condition to false but not the entire condition

```
sched : prev_pid != 0
sched : !(prev_pid == 0)
```

evaluates to:

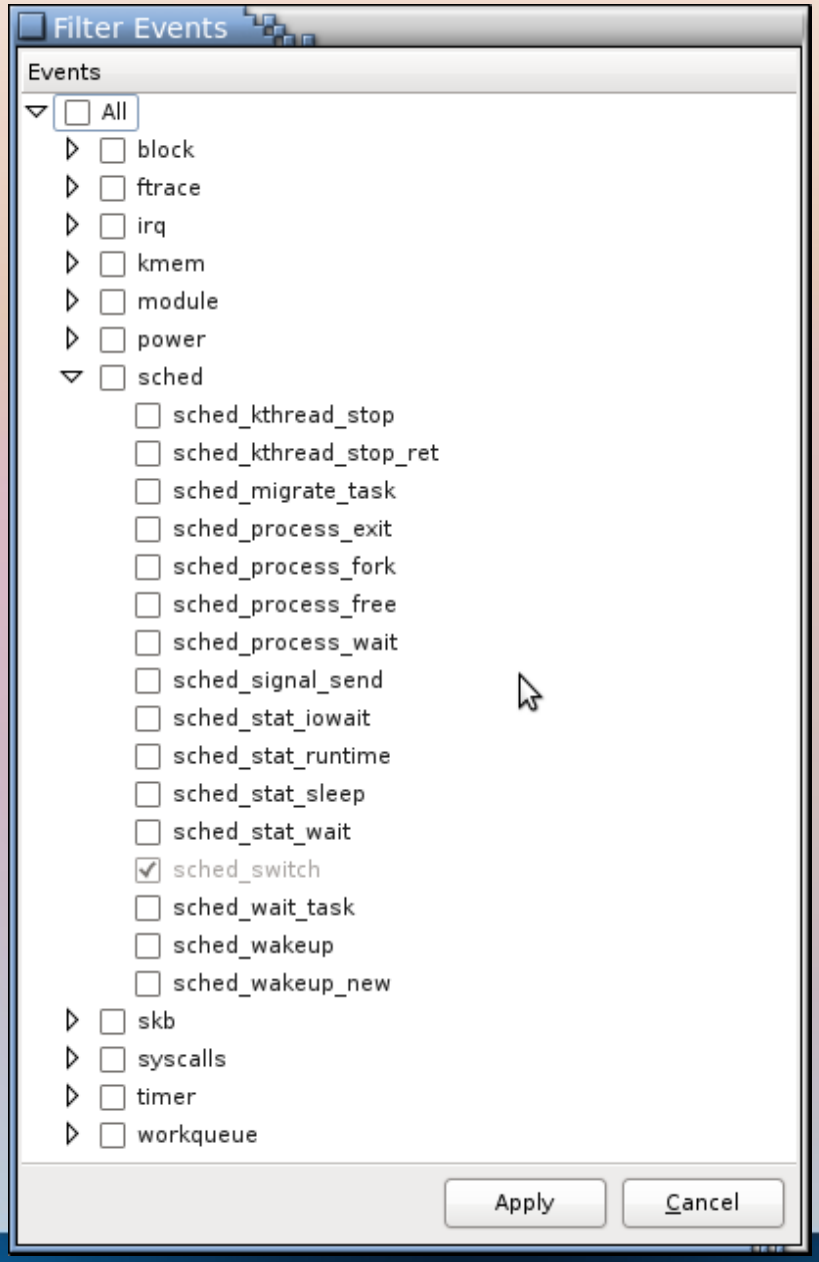
```
sched : FALSE
sched : !(FALSE)
```

Comparing Strings

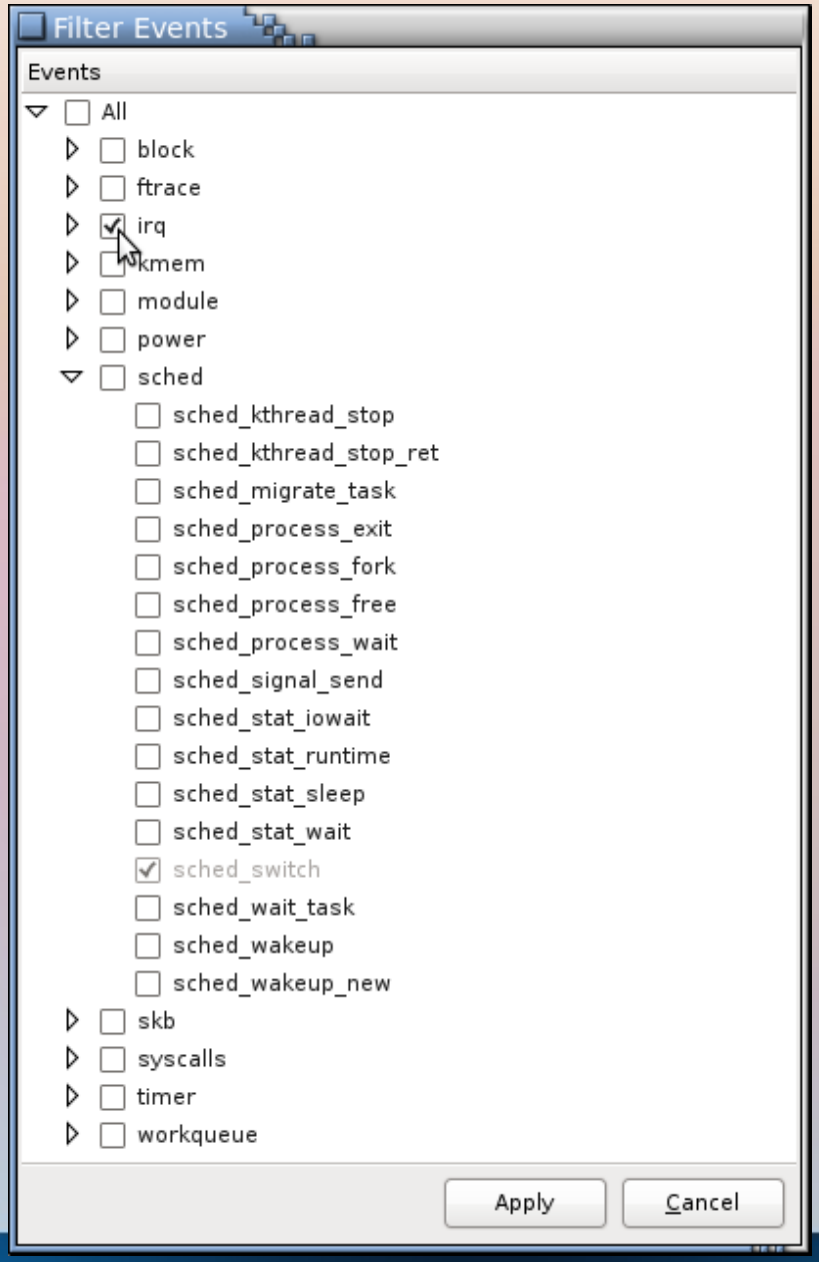
- Strings can compare with regular expressions
  - regex(7)
  - Use =~ or !~

```
sched_switch : next_comm =~ "^events/[23]$"
```

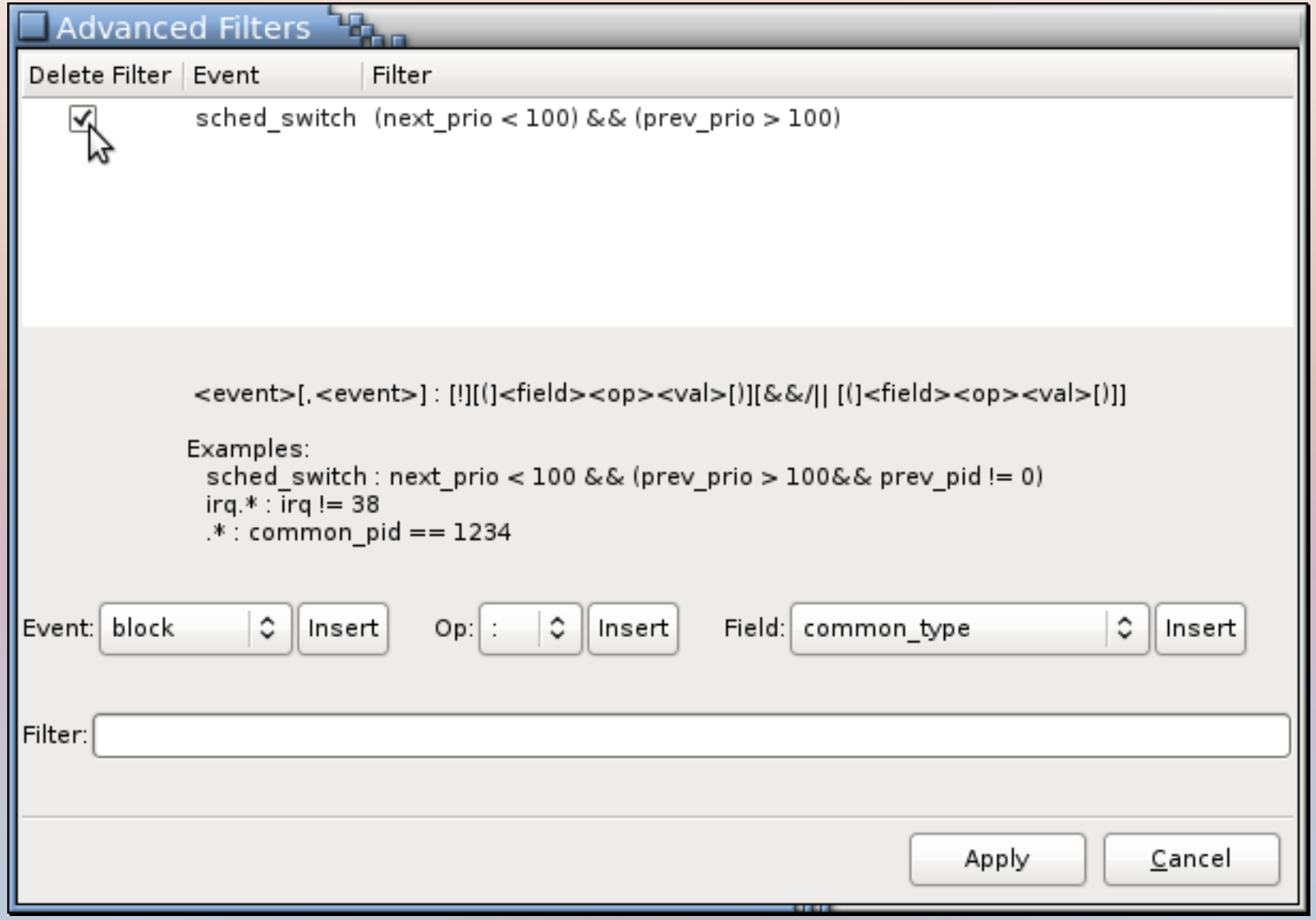
Event Filters with Advanced



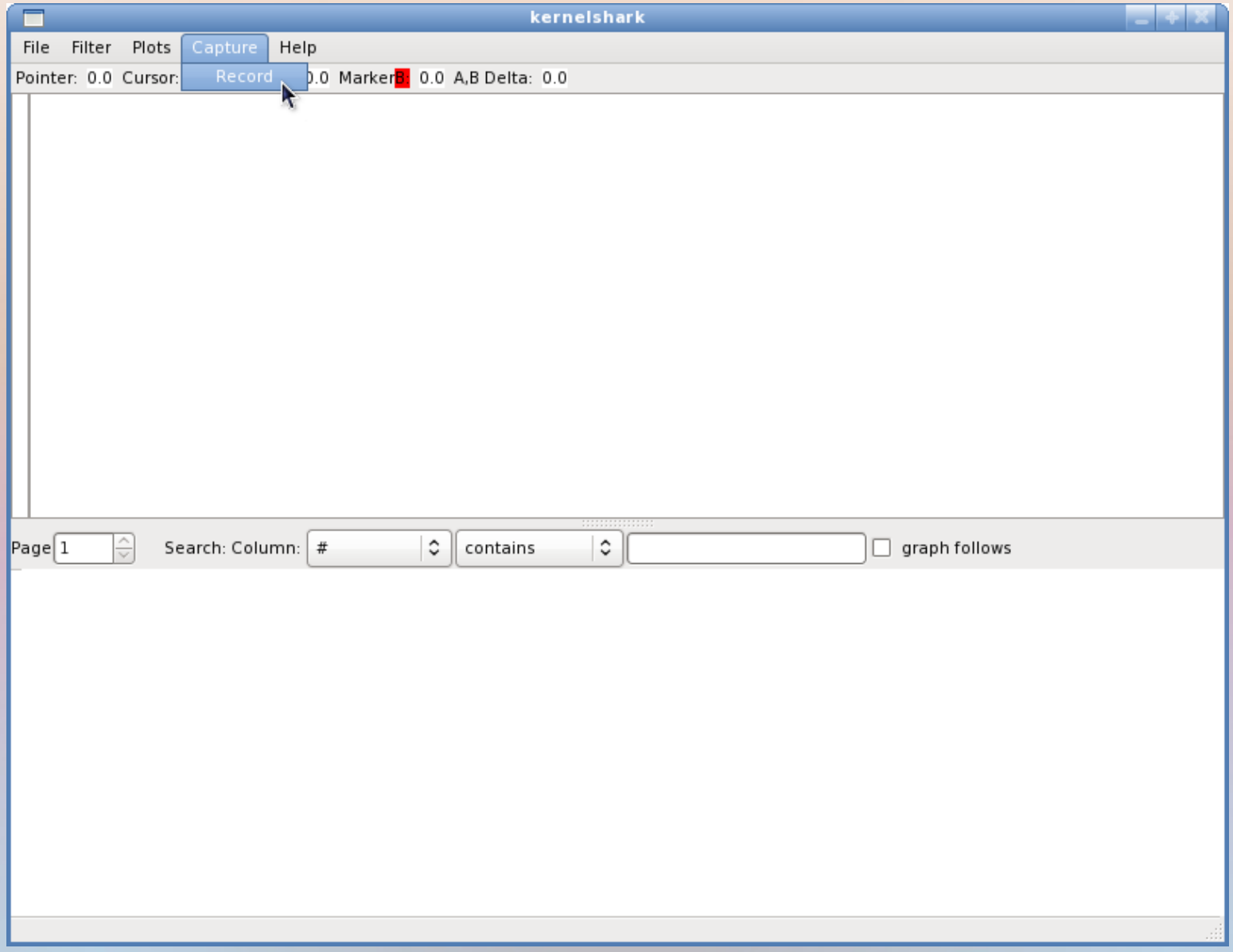
Adding Events after Advance



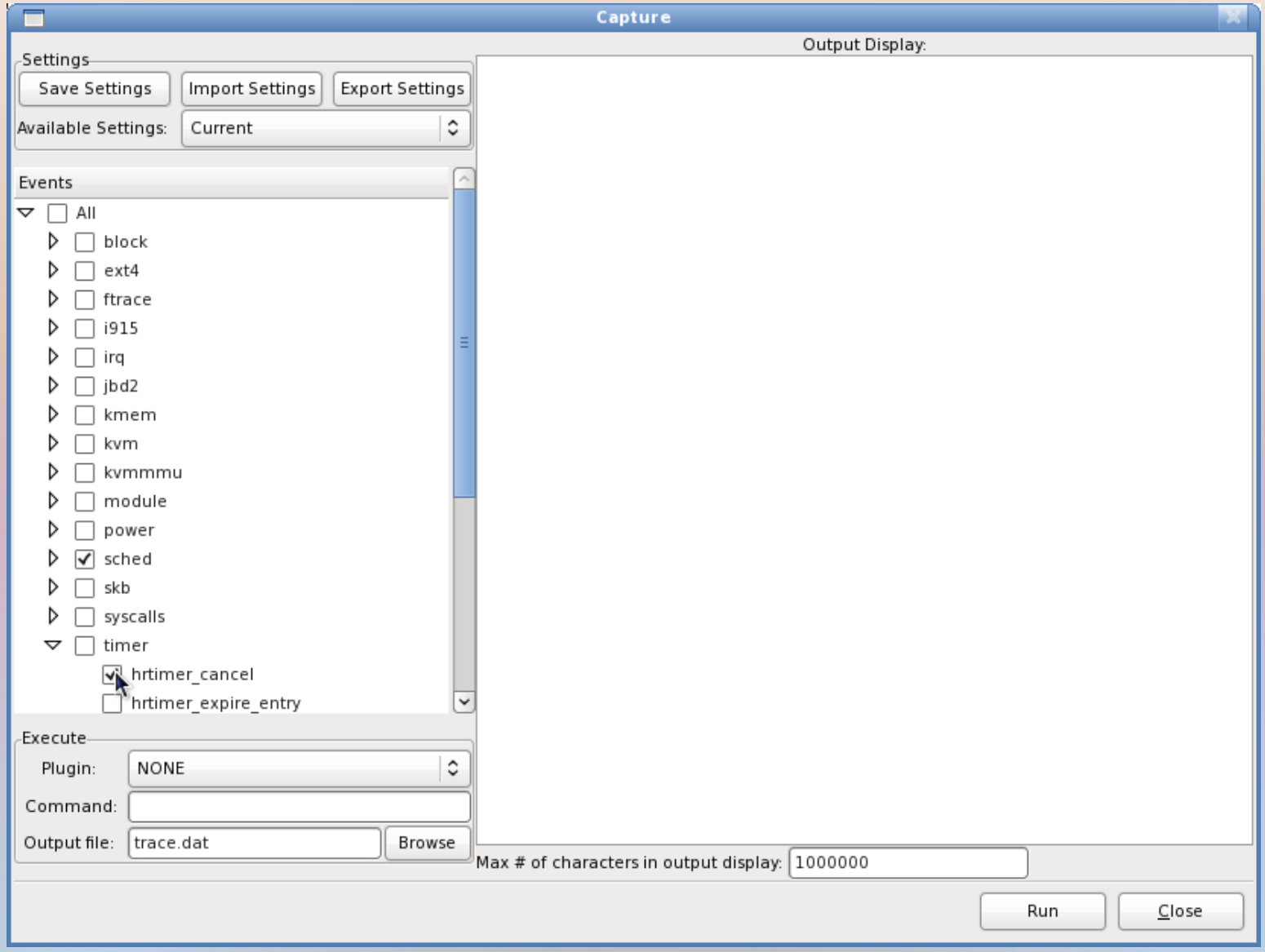
Deleting Advanced Filters



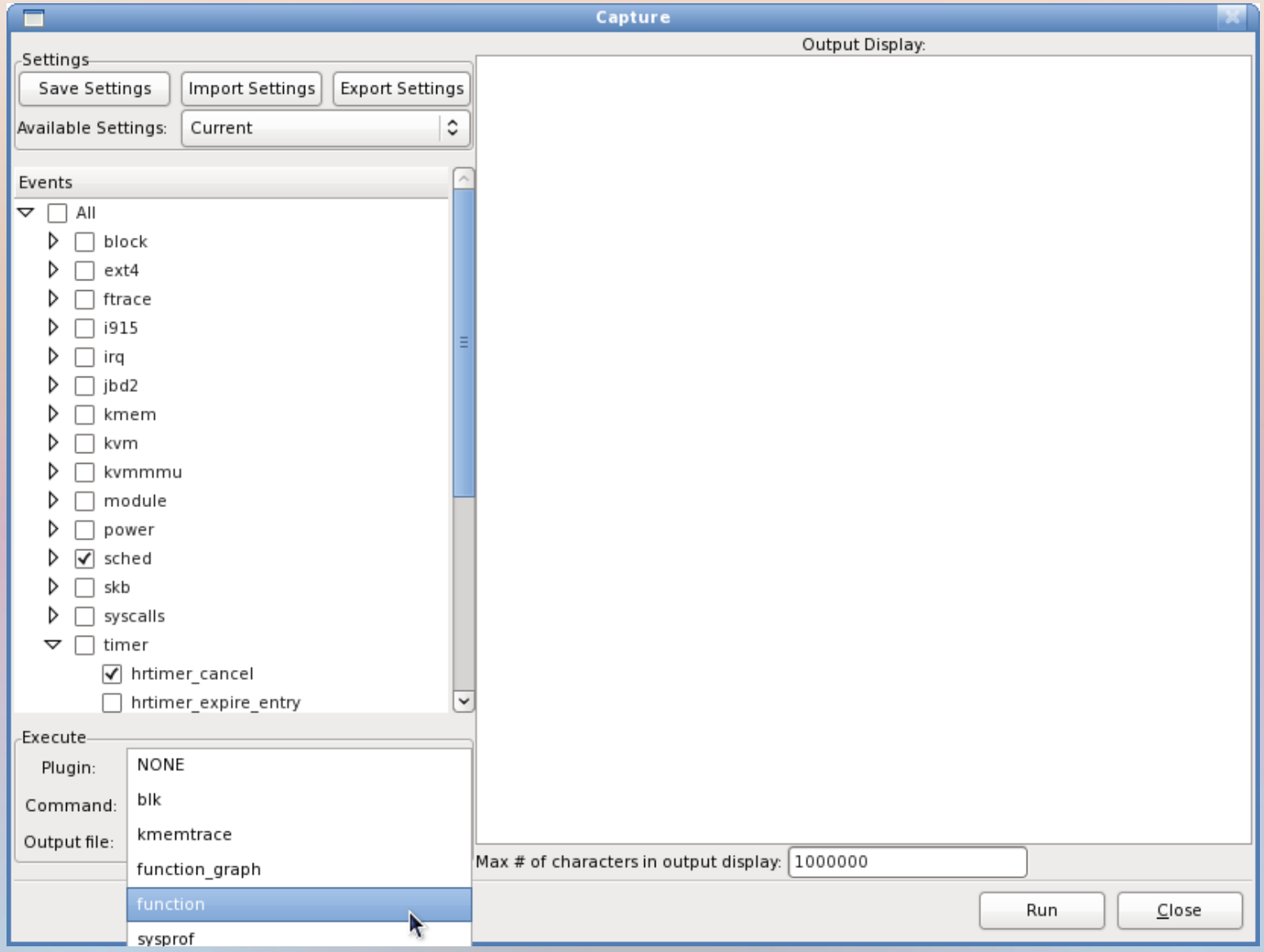
Recording


kernelshark
File
Filter
Plots
Capture
Help
Pointer: 0.0 Cursor:
Record
0.0 MarkerB: 0.0 A,B Delta: 0.0
Page 1
Search: Column: #
contains
graph follows

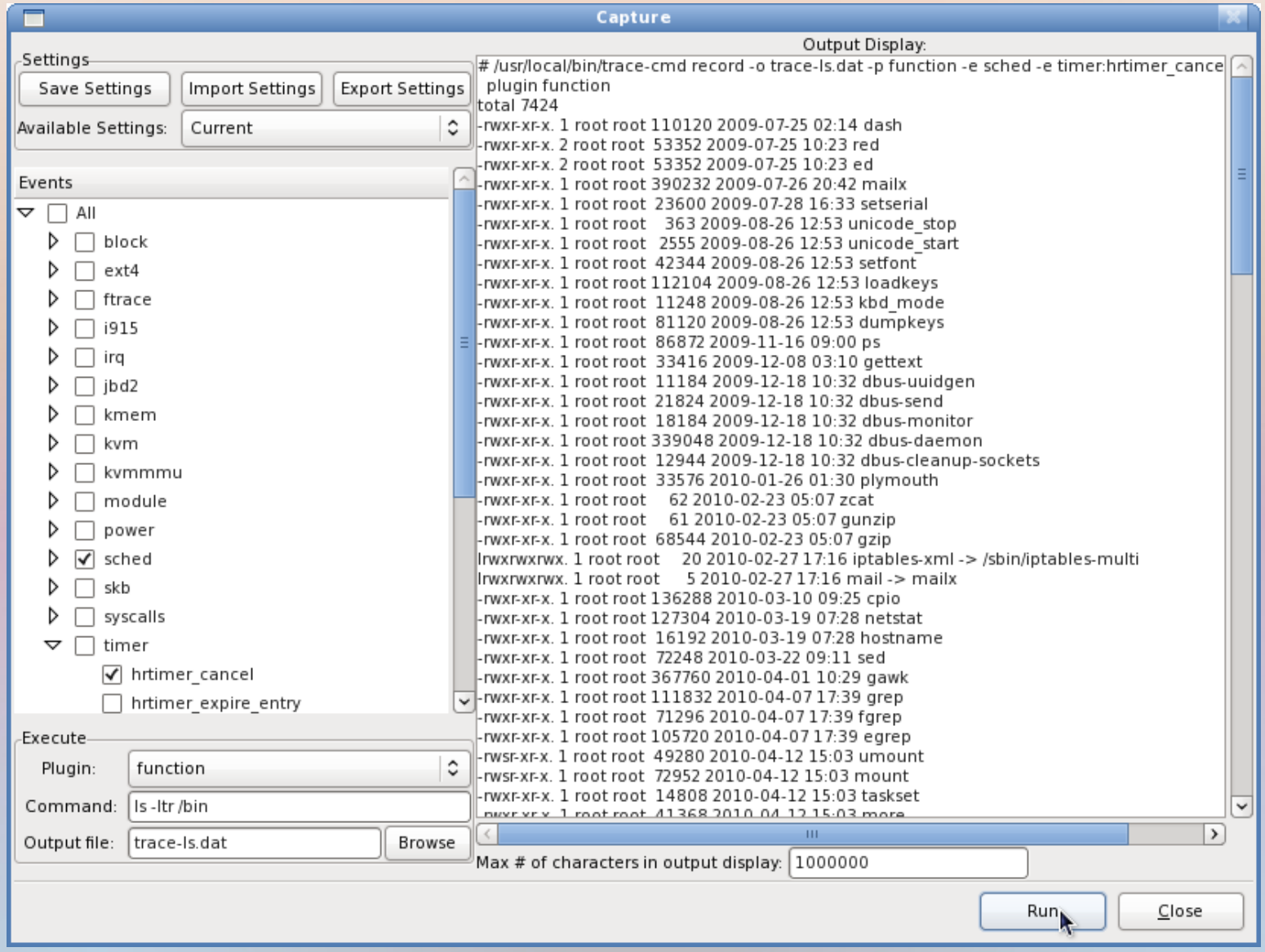
Recording


Capture
Settings
Save Settings
Import Settings
Export Settings
Available Settings:
Current
Events
All
block
ext4
ftrace
i915
irq
jbd2
kmem
kvm
kvmmmu
module
power
sched
skb
syscalls
timer
hrtimer_cancel
hrtimer_expire_entry
Execute
Plugin:
NONE
Command:
Output file:
trace.dat
Browse
Output Display:
Max # of characters in output display:
1000000
Run
Close

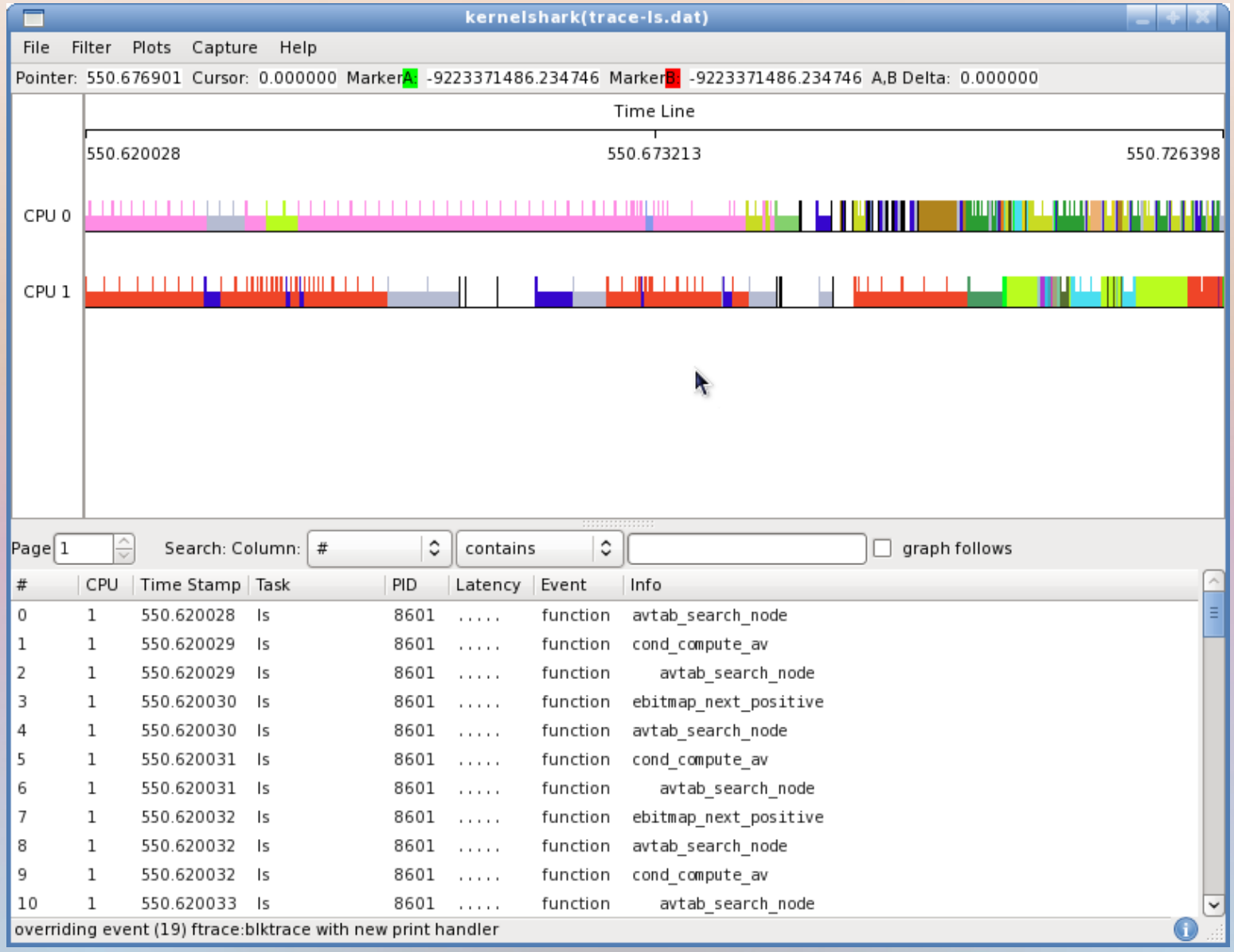
Recording


Capture
Settings
Save Settings
Import Settings
Export Settings
Available Settings:
Current
Events
All
block
ext4
ftrace
i915
irq
jbd2
kmem
kvm
kvmmmu
module
power
sched
skb
syscalls
timer
hrtimer_cancel
hrtimer_expire_entry
Execute
Plugin:
Command:
Output file:
NONE
blk
kmemtrace
function_graph
function
sysprof
Output Display:
Max # of characters in output display:
1000000
Run
Close

Recording


Capture
Settings
Save Settings
Import Settings
Export Settings
Available Settings:
Current
Events
All
block
ext4
ftrace
i915
irq
jbd2
kmem
kvm
kvmmmu
module
power
sched
skb
syscalls
timer
hrtimer_cancel
hrtimer_expire_entry
Execute
Plugin:
function
Command:
ls -ltr /bin
Output file:
trace-ls.dat
Browse
Output Display:
/usr/local/bin/trace-cmd record -o trace-ls.dat -p function -e sched -e timer:hrtimer_cance
plugin function
total 7424
-rwxr-xr-x. 1 root root 110120 2009-07-25 02:14 dash
-rwxr-xr-x. 2 root root 53352 2009-07-25 10:23 red
-rwxr-xr-x. 2 root root 53352 2009-07-25 10:23 ed
-rwxr-xr-x. 1 root root 390232 2009-07-26 20:42 mailx
-rwxr-xr-x. 1 root root 23600 2009-07-28 16:33 setserial
-rwxr-xr-x. 1 root root 363 2009-08-26 12:53 unicode_stop
-rwxr-xr-x. 1 root root 2555 2009-08-26 12:53 unicode_start
-rwxr-xr-x. 1 root root 42344 2009-08-26 12:53 setfont
-rwxr-xr-x. 1 root root 112104 2009-08-26 12:53 loadkeys
-rwxr-xr-x. 1 root root 11248 2009-08-26 12:53 kbd_mode
-rwxr-xr-x. 1 root root 81120 2009-08-26 12:53 dumpkeys
-rwxr-xr-x. 1 root root 86872 2009-11-16 09:00 ps
-rwxr-xr-x. 1 root root 33416 2009-12-08 03:10 gettext
-rwxr-xr-x. 1 root root 11184 2009-12-18 10:32 dbus-uuidgen
-rwxr-xr-x. 1 root root 21824 2009-12-18 10:32 dbus-send
-rwxr-xr-x. 1 root root 18184 2009-12-18 10:32 dbus-monitor
-rwxr-xr-x. 1 root root 339048 2009-12-18 10:32 dbus-daemon
-rwxr-xr-x. 1 root root 12944 2009-12-18 10:32 dbus-cleanup-sockets
-rwxr-xr-x. 1 root root 33576 2010-01-26 01:30 plymouth
-rwxr-xr-x. 1 root root 62 2010-02-23 05:07 zcat
-rwxr-xr-x. 1 root root 61 2010-02-23 05:07 gunzip
-rwxr-xr-x. 1 root root 68544 2010-02-23 05:07 gzip
lrwxrwxrwx. 1 root root 20 2010-02-27 17:16 iptables-xml -> /sbin/iptables-multi
lrwxrwxrwx. 1 root root 5 2010-02-27 17:16 mail -> mailx
-rwxr-xr-x. 1 root root 136288 2010-03-10 09:25 cpio
-rwxr-xr-x. 1 root root 127304 2010-03-19 07:28 netstat
-rwxr-xr-x. 1 root root 16192 2010-03-19 07:28 hostname
-rwxr-xr-x. 1 root root 72248 2010-03-22 09:11 sed
-rwxr-xr-x. 1 root root 367760 2010-04-01 10:29 gawk
-rwxr-xr-x. 1 root root 111832 2010-04-07 17:39 grep
-rwxr-xr-x. 1 root root 71296 2010-04-07 17:39 fgrep
-rwxr-xr-x. 1 root root 105720 2010-04-07 17:39 egrep
-rwsr-xr-x. 1 root root 49280 2010-04-12 15:03 umount
-rwsr-xr-x. 1 root root 72952 2010-04-12 15:03 mount
-rwxr-xr-x. 1 root root 14808 2010-04-12 15:03 taskset
Max # of characters in output display:
1000000
Run
Close

Recording


kernelshark(trace-ls.dat)
File Filter Plots Capture Help
Pointer: 550.676901 Cursor: 0.000000 MarkerA: -9223371486.234746 MarkerB: -9223371486.234746 A,B Delta: 0.000000
Time Line
550.620028
550.673213
550.726398
CPU 0
CPU 1
Page 1 Search: Column: # contains graph follows
CPU Time Stamp Task PID Latency Event Info
0 1 550.620028 ls 8601 ..... function avtab_search_node
1 1 550.620029 ls 8601 ..... function cond_compute_av
2 1 550.620029 ls 8601 ..... function avtab_search_node
3 1 550.620030 ls 8601 ..... function ebitmap_next_positive
4 1 550.620030 ls 8601 ..... function avtab_search_node
5 1 550.620031 ls 8601 ..... function cond_compute_av
6 1 550.620031 ls 8601 ..... function avtab_search_node
7 1 550.620032 ls 8601 ..... function ebitmap_next_positive
8 1 550.620032 ls 8601 ..... function avtab_search_node
9 1 550.620032 ls 8601 ..... function cond_compute_av
10 1 550.620033 ls 8601 ..... function avtab_search_node
overriding event (19) ftrace:blktrace with new print handler

Kernel Shark

Demo!

Questions?