Ubuntu Developer Summit 2010

Squashing system-wide bugs with LTTng

E-mail:
mathieu.desnoyers@efficios.com

> Presenter

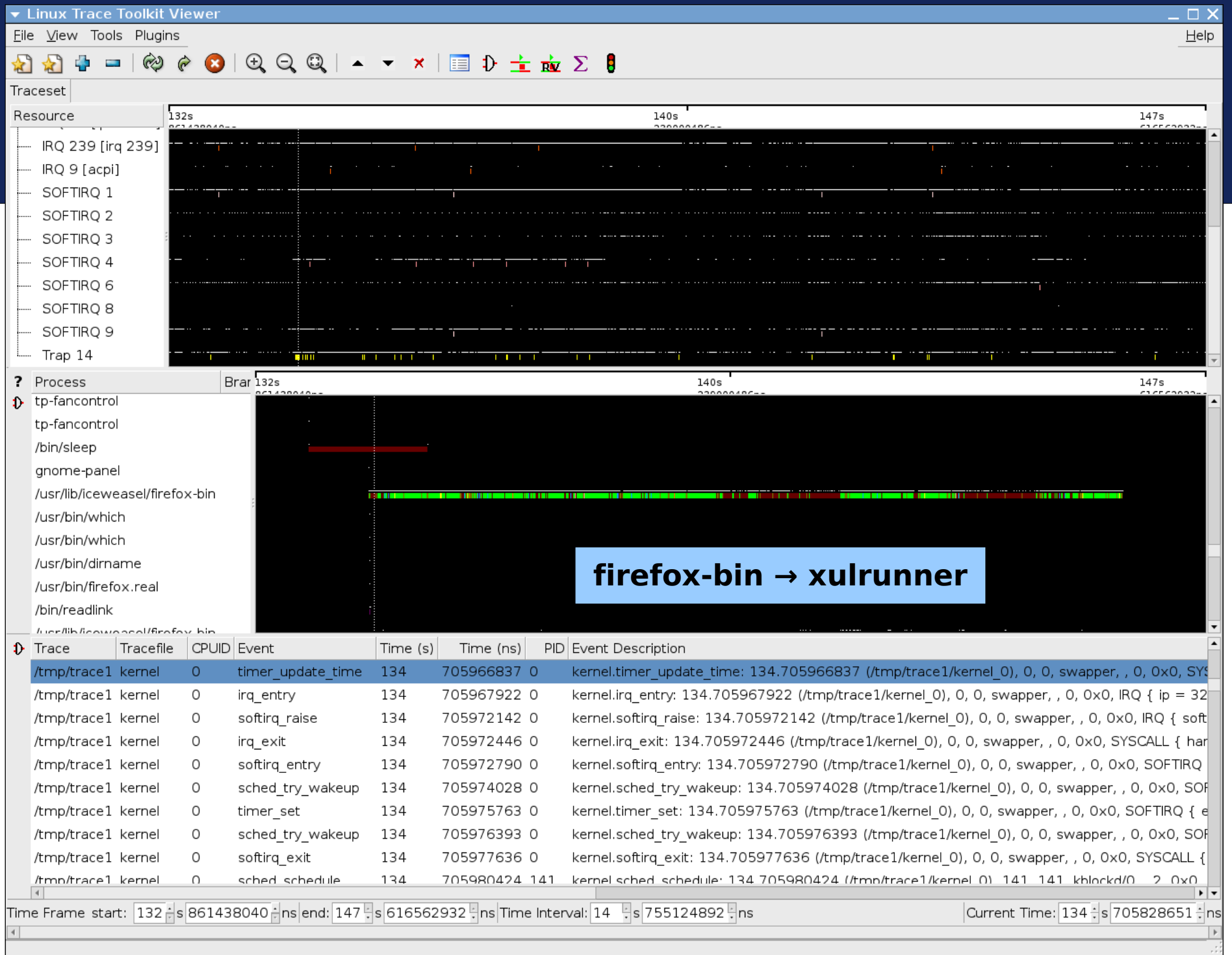
- Mathieu Desnoyers
- EfficiOS Inc.
  - http://www.efficios.com
- Author/Maintainer of
  - LTTng, LTTV, Userspace RCU

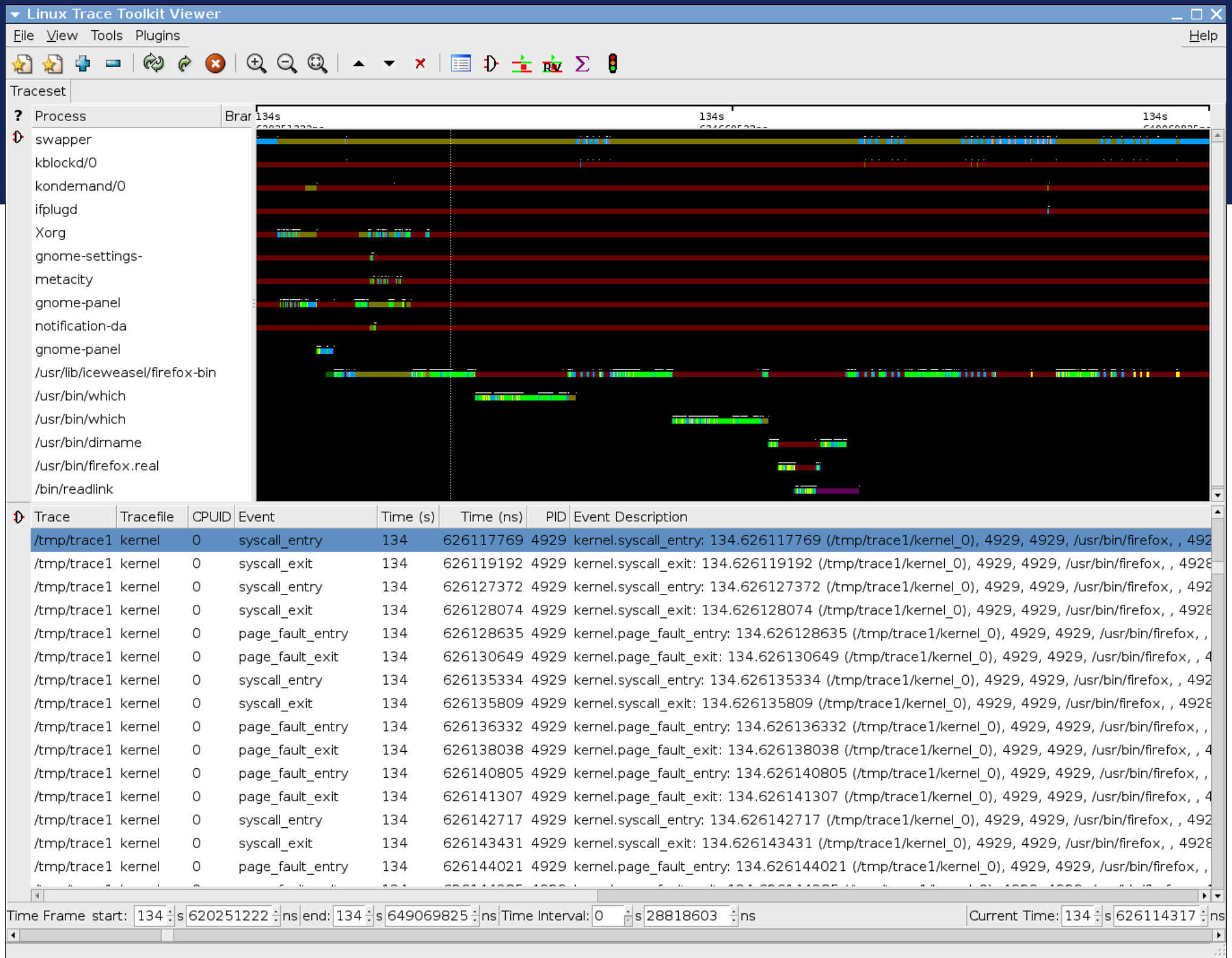
What is LTTng ?

- Linux Trace Toolkit Next Generation
- LTTng is a
  - Kernel tracer
  - Userland application/library tracer (UST)
- Provides synchronized kernel-userspace traces
- Very low performance impact

> How is tracing useful to you ?

- Developer
  - Narrow down bug causes
  - Identify performance hogs
- End user
  - Provide more detailed application and kernel crash reports
  - Could be automated if integrated with distribution


Linux Trace Toolkit Viewer
File View Tools Plugins
Help
Traceset
Resource
132s
140s
147s
IRQ 239 [irq 239]
IRQ 9 [acpi]
SOFTIRQ 1
SOFTIRQ 2
SOFTIRQ 3
SOFTIRQ 4
SOFTIRQ 6
SOFTIRQ 8
SOFTIRQ 9
Trap 14
Process
tp-fancontrol
tp-fancontrol
/bin/sleep
gnome-panel
/usr/lib/iceweasel/firefox-bin
/usr/bin/which
/usr/bin/which
/usr/bin/dirname
/usr/bin/firefox.real
/bin/readlink
firefox-bin → xulrunner
Trace Tracefile CPUID Event Time (s) Time (ns) PID Event Description
/tmp/trace1 kernel 0 timer_update_time 134 705966837 0 kernel.timer_update_time: 134.705966837 (/tmp/trace1/kernel_0), 0, 0, swapper, , 0, 0x0, SY
/tmp/trace1 kernel 0 irq_entry 134 705967922 0 kernel.irq_entry: 134.705967922 (/tmp/trace1/kernel_0), 0, 0, swapper, , 0, 0x0, IRQ { ip = 32
/tmp/trace1 kernel 0 softirq_raise 134 705972142 0 kernel.softirq_raise: 134.705972142 (/tmp/trace1/kernel_0), 0, 0, swapper, , 0, 0x0, IRQ { soft
/tmp/trace1 kernel 0 irq_exit 134 705972446 0 kernel.irq_exit: 134.705972446 (/tmp/trace1/kernel_0), 0, 0, swapper, , 0, 0x0, SYSCALL { har
/tmp/trace1 kernel 0 softirq_entry 134 705972790 0 kernel.softirq_entry: 134.705972790 (/tmp/trace1/kernel_0), 0, 0, swapper, , 0, 0x0, SOFTIRQ
/tmp/trace1 kernel 0 sched_try_wakeup 134 705974028 0 kernel.sched_try_wakeup: 134.705974028 (/tmp/trace1/kernel_0), 0, 0, swapper, , 0, 0x0, SOF
/tmp/trace1 kernel 0 timer_set 134 705975763 0 kernel.timer_set: 134.705975763 (/tmp/trace1/kernel_0), 0, 0, swapper, , 0, 0x0, SOFTIRQ { e
/tmp/trace1 kernel 0 sched_try_wakeup 134 705976393 0 kernel.sched_try_wakeup: 134.705976393 (/tmp/trace1/kernel_0), 0, 0, swapper, , 0, 0x0, SOF
/tmp/trace1 kernel 0 softirq_exit 134 705977636 0 kernel.softirq_exit: 134.705977636 (/tmp/trace1/kernel_0), 0, 0, swapper, , 0, 0x0, SYSCALL {
/tmp/trace1 kernel 0 sched_schedule 134 705980424 141 kernel.sched_schedule: 134.705980424 (/tmp/trace1/kernel_0), 141, 141, kblockd/0, 2, 0x0
Time Frame start: 132 s 861438040 ns end: 147 s 616562932 ns Time Interval: 14 s 755124892 ns
Current Time: 134 s 705828651 ns


Linux Trace Toolkit Viewer
File View Tools Plugins
Help
Traceset
Process
Brar
134s
134s
134s
swapper
kblockd/0
kondemand/0
ifplugd
Xorg
gnome-settings-
metacity
gnome-panel
notification-da
gnome-panel
/usr/lib/iceweasel/firefox-bin
/usr/bin/which
/usr/bin/which
/usr/bin/dirname
/usr/bin/firefox.real
/bin/readlink
Trace Tracefile CPUID Event Time (s) Time (ns) PID Event Description
/tmp/trace1 kernel 0 syscall_entry 134 626117769 4929 kernel.syscall_entry: 134.626117769 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, , 492
/tmp/trace1 kernel 0 syscall_exit 134 626119192 4929 kernel.syscall_exit: 134.626119192 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, , 4928
/tmp/trace1 kernel 0 syscall_entry 134 626127372 4929 kernel.syscall_entry: 134.626127372 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, , 492
/tmp/trace1 kernel 0 syscall_exit 134 626128074 4929 kernel.syscall_exit: 134.626128074 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, , 4928
/tmp/trace1 kernel 0 page_fault_entry 134 626128635 4929 kernel.page_fault_entry: 134.626128635 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, ,
/tmp/trace1 kernel 0 page_fault_exit 134 626130649 4929 kernel.page_fault_exit: 134.626130649 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, , 4
/tmp/trace1 kernel 0 syscall_entry 134 626135334 4929 kernel.syscall_entry: 134.626135334 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, , 492
/tmp/trace1 kernel 0 syscall_exit 134 626135809 4929 kernel.syscall_exit: 134.626135809 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, , 4928
/tmp/trace1 kernel 0 page_fault_entry 134 626136332 4929 kernel.page_fault_entry: 134.626136332 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, ,
/tmp/trace1 kernel 0 page_fault_exit 134 626138038 4929 kernel.page_fault_exit: 134.626138038 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, , 4
/tmp/trace1 kernel 0 page_fault_entry 134 626140805 4929 kernel.page_fault_entry: 134.626140805 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, ,
/tmp/trace1 kernel 0 page_fault_exit 134 626141307 4929 kernel.page_fault_exit: 134.626141307 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, , 4
/tmp/trace1 kernel 0 syscall_entry 134 626142717 4929 kernel.syscall_entry: 134.626142717 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, , 492
/tmp/trace1 kernel 0 syscall_exit 134 626143431 4929 kernel.syscall_exit: 134.626143431 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, , 4928
/tmp/trace1 kernel 0 page_fault_entry 134 626144021 4929 kernel.page_fault_entry: 134.626144021 (/tmp/trace1/kernel_0), 4929, 4929, /usr/bin/firefox, ,
Time Frame start: 134 s 620251222 ns end: 134 s 649069825 ns Time Interval: 0 s 28818603 ns
Current Time: 134 s 626114317 ns

> Text-based mode, filtering example

compudj@compumobile: ~/local/bin

File Edit View Terminal Help

```
compudj@compumobile:~/local/bin$ ./lttv -m textDump -e "state.pid=4929&event.name=fs.open" -t /traces/trace1 |head -n 15
Trace set contains 1 traces

fs.open: 134.622638711 (/traces/trace1/fs_0), 4929, 4929, gnome-panel, , 4928, 0x0, SYSCALL { fd = 29, filename = "/proc/
self/fd" }
fs.open: 134.622783466 (/traces/trace1/fs_0), 4929, 4929, gnome-panel, , 4928, 0x0, SYSCALL { fd = 29, filename = "/dev/n
ull" }
fs.open: 134.625092483 (/traces/trace1/fs_0), 4929, 4929, /usr/bin/firefox, , 4928, 0x0, SYSCALL { fd = 3, filename = "/e
tc/ld.so.cache" }
fs.open: 134.625120094 (/traces/trace1/fs_0), 4929, 4929, /usr/bin/firefox, , 4928, 0x0, SYSCALL { fd = 3, filename = "/l
ib/libncurses.so.5" }
fs.open: 134.625168232 (/traces/trace1/fs_0), 4929, 4929, /usr/bin/firefox, , 4928, 0x0, SYSCALL { fd = 3, filename = "/l
ib/i686/cmov/libdl.so.2" }
fs.open: 134.625203622 (/traces/trace1/fs_0), 4929, 4929, /usr/bin/firefox, , 4928, 0x0, SYSCALL { fd = 3, filename = "/l
ib/i686/cmov/libc.so.6" }
fs.open: 134.625628479 (/traces/trace1/fs_0), 4929, 4929, /usr/bin/firefox, , 4928, 0x0, SYSCALL { fd = -6, filename = "/
dev/tty" }
fs.open: 134.625684640 (/traces/trace1/fs_0), 4929, 4929, /usr/bin/firefox, , 4928, 0x0, SYSCALL { fd = 3, filename = "/u
sr/lib/locale/locale-archive" }
fs.open: 134.625740732 (/traces/trace1/fs_0), 4929, 4929, /usr/bin/firefox, , 4928, 0x0, SYSCALL { fd = 3, filename = "/u
sr/lib/locale/locale-archive" }
fs.open: 134.625835844 (/traces/trace1/fs_0), 4929, 4929, /usr/bin/firefox, , 4928, 0x0, SYSCALL { fd = 3, filename = "/p
roc/meminfo" }
fs.open: 134.626053763 (/traces/trace1/fs_0), 4929, 4929, /usr/bin/firefox, , 4928, 0x0, SYSCALL { fd = 3, filename = "/u
sr/lib/gconv/gconv-modules.cache" }
fs.open: 134.626650096 (/traces/trace1/fs_0), 4929, 4929, /usr/bin/firefox, , 4928, 0x0, SYSCALL { fd = 3, filename = "/u
sr/bin/firefox" }
fs.open: 134.631100446 (/traces/trace1/fs_0), 4929, 4929, /usr/bin/firefox.real, , 4928, 0x0, SYSCALL { fd = 3, filename
= "/etc/ld.so.cache" }
compudj@compumobile:~/local/bin$
```

> Typical tracing workflow

- Given a reproducible problem
- Gather trace (high level at first)
- Analyze (narrow down the problem source)
- Add instrumentation (e.g. anchor identifying the problem) if needed
- Rince, repeat

> War stories

- HAL driver debugging
- Linux scheduler latency identification
- Application slowdown due to distributed file system cache size issues
- Identify callers of "sync()" and "fdatasync()"
- Identify bugs in the kernel VM

> Availability

- Ubuntu: PPA packages available, for both kernel and user-space tracing
  - https://launchpad.net/~lttng/+archive/ppa
  - http://lttng.org
- Embedded distributions (WindRiver, STLinux, Montavista, ...)
- Projects using UST
  - Qemu/KVM is currently instrumented with UST
  - MariaDB instrumentation (upcoming)

> Questions ?

EfficiOS

- http://www.efficios.com

- LTTng Information
  - http://lttng.org
  - ltt-dev@lists.casi.polymtl.ca