
Subject: Re: [PATCH 0/7] Profiling sleep times (v3)
Posted by [Arun Sharma](#) on Wed, 07 Dec 2011 20:33:48 GMT
[View Forum Message](#) <> [Reply to Message](#)

On 12/5/11 11:15 PM, Andrey Vagin wrote:

> Arun Sharma said, that the second versions of patches works ok for him.
> (Arun is the first user of this functionality after me.)

Yes - Andrey's patches (v2) have been functional for me when used via:

```
perf record -agP -e sched:sched_switch --filter "prev_state == 1 ||  
prev_state == 2" -e sched:sched_stat_sleep,sched:sched_stat_iowait --  
sleep 3  
mv perf.data{,.old}; perf inject -s -i perf.data.old -o perf.data  
perf report --stdio -g graph --sort pid -C command-name
```

There are two major issues though:

* The above command lines cause us to collect way too much data and perf
can't keep up on a busy server

Eg:

Warning:
Processed 55182 events and lost 13 chunks!

Check IO/CPU overload!

* Requires root access

I suspect there is a way to stash the delay information available at
enqueue_sleeper() into the task_struct (or some place else) and make it
available at the sched:sched_switch tracepoint. This would solve both of
the problems above, but I don't have a patch yet.

-Arun
