Subject: Re: [PATCH v3 5/6] Also record sleep start for a task group
Posted by Glauber Costa on Wed, 30 May 2012 12:44:09 GMT

On 05/30/2012 04:44 PM, Peter Zijlstra wrote:
> On Wed, 2012-05-30 at 16:24 +0400, Glauber Costa wrote:
>> sleep_start is not for iowait. This is for idle. And I know no other way
>> to collect idle time per cgroup, other than the time during which it was
>> out of the runqueue.
>>
>> Now what you say about the sleepers don't make that much sense for idle
>> because this information is per-cpu as well.
>>
>> When the se is being dequeued, it means none of its children is running
>> on that runqueue. That's idle.
>
> But does that mean the cgroup is idle? Its impossible to re-construct
> the machine state from this per-cpu data if your definition of
> cgroup-idle is the time when _all_ cpus are idle.
>
It is idle for that runqueue, aka cpu. The cgroup itself is idle when
all cpus are idle. And yes, then you have 2s per sec of idle in a 2-way
system.

That's pretty much how a physical box works as well.