
Subject: Re: [PATCH 2/2] CFS CGroup: Report usage
Posted by [Srivatsa Vaddagiri](#) on Wed, 24 Oct 2007 04:29:37 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Tue, Oct 23, 2007 at 07:28:22PM -0700, Paul Menage wrote:

> > Good point. I think we need to subtract out the time it was waiting on runqueue
> > when calculating idle time.

> >

> > |----- -----zzzzzzzzzzzz.....-----|
> > t0 t1 t2 t3 t4 t5 t6

> >

> >

> > ---- -> Running time

> > -> Waiting time (to get on the cpu)

> > zzzz -> Sleeping time (when it didnt want to run because of
> > lack of tasks)

> >

> > So, in this case,

> >

> > idle time = (t4 - t3) / [(t6 - t1) - (t2-t1) - (t5-t4)

> >

>

> Do you mean (t6 - t0) where you have (t6 - t1)?

Ah ..yes.

> > ?

> >

> > This idle time will be a per-cpu stat for every cgroup and needs to be
> > consolidated across cpus into a single idle-stat number, just like how
> > top does it.

>

> This would be an idle fraction, not an idle time. (seconds divided by seconds)

agreed, we need to be reporting idle time in (milli)seconds, although
the requirement we had was to report it back in percentage. I guess the
percentage figure can be derived from the raw idle time number.

How about:

idle time = t4-t3 (effectively sleep time)

in the above example?

> It doesn't seem quite right to me that a cgroup's idle time metric be
> affected by the activity of other cgroups on the machine,

I don't see how the idle time metric defined above (t4-t3) can be

affected by other cgroup activity, unless the execution pattern of one cgroup is dependent on the others.

However the minute you translate this idle time into a percentage wrt wall-clock time, then yes a cgroup's idle percentage can be affected by others. For idle percentage to be meaningful, I would imagine that user-space tools will need to calculate it after discarding a group's wait time.

- > but it's
- > hard to come up with a way of measuring it that doesn't have this
- > behaviour - which is why, in the absence of hard CPU partitioning,
- > it's not clear to me how much use this would be.
- >
- > What would people be planning to use it for?

I think primarily for systems management tools to report back various statistics about a OpenVZ/VServer-like container (just like top reports stats for a OS currently). Let me check if there are other uses envisioned for it.

--

Regards,
vatsa

Containers mailing list
Containers@lists.linux-foundation.org
<https://lists.linux-foundation.org/mailman/listinfo/containers>
