Subject: Re: Revert for cgroups CPU accounting subsystem patch
Posted by Srivatsa Vaddagiri on Tue, 13 Nov 2007 07:37:40 GMT
View Forum Message <> Reply to Message

On Mon, Nov 12, 2007 at 10:05:24PM -0800, Paul Menage wrote:
> On Nov 12, 2007 10:00 PM, Srivatsa Vaddagiri <vatsa@linux.vnet.ibm.com> wrote:
> > On second thoughts, this may be a usefull controller of its own.
> > Say I just want to "monitor" usage (for accounting purpose) of a group of
> > tasks, but don't want to control their cpu consumption, then cpuacct
> > controller would come in handy.
> >
>
> That's plausible, but having two separate ways of tracking and
> reporting the CPU usage of a cgroup seems wrong.
>
> How bad would it be in your suggested case if you just give each
> cgroup the same weight?

That's still introducing a deviation from the normal behavior we would
have had we allowed all tasks to be part of the same "control" group/runqueue.

For ex: using nice value to vary bandwidth between tasks makes sense if
they are all part of the same group.

Also an application with more tasks will get more cpu power (as intended)
compared to another app with less tasks, provided they are all
part of the same group.

Regarding your concern about tracking cpu usage in different ways, it
could be mitigated if we have cpuacct controller track usage as per
information present in a task's sched entity structure
(tsk->se.sum_exec_runtime) i.e call cpuacct_charge() from
__update_curr() which would accumulate the execution time of the
group in a SMP friendly manner (i.e dump it in a per-cpu per-group counter
first and then aggregate to a global per-group counter).

This will let account and control grouping to be independent if desired.

What do you think?

> So there would be fair scheduling between
> cgroups, which seems as reasonable as any other choice in the event
> that the CPU is contended.

--
Regards,
vatsa

_____

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