Subject: Re: [PATCH] task containersv11 add tasks file interface fix for cpusets
Posted by Paul Jackson on Wed, 03 Oct 2007 20:52:32 GMT
View Forum Message <> Reply to Message

> What was wrong with my suggestion from a couple of emails back? Adding
> the following in cpuset_attach():
>
> struct cgroup_iter it;
> struct task_struct *p;
> while ((p = cgroup_iter_next(cs->css.cgroup, &it))) {
> set_cpus_allowed(p, cs->cpus_allowed);
> }
> cgroup_iter_end(cs->css.cgroup, &it);

Hmmm ... that just might work.

And this brings to light the reason (justification, excuse, whatever
you call it) that I probably didn't do this earlier.

In the dark ages before cgroups (aka containers) we did not have an
efficient way to walk the tasks in a cpuset. One had to walk the entire
task list, comparing task struct cpuset pointers. On big honking NUMA
iron, one should avoid task list walks as much as one can get away
with, even if it meant sneaking in a little bit racey API. Since some
updates of a cpusets 'cpus' mask don't need it (you happen to know that
all tasks in that cpuset are pause'd anyway) I might have made the
tradeoff to make this task list walk an explicitly invoked user action,
to be done only when needed.

But now (correct me if I'm wrong here) cgroups has a per-cgroup task
list, and the above loop has cost linear in the number of tasks
actually in the cgroup, plus (unfortunate but necessary and tolerable)
the cost of taking a global css_set_lock, right?

And I take it the above code snipped is missing the cgroup_iter_start,
correct?

I'll ask Andrew to kill this patch of mine, and I will test out your
suggestion this evening.

This still leaves the other creepy crawlies involving cpusets and hot
plug that I glimpsed slithering by in my last message. I guess I'll
start a separate discussion with Cliff Wickman and whomever else I think
might want to be involved on those issues.

Nice work - thanks.

--

I won't rest till it's the best ...
Programmer, Linux Scalability
Paul Jackson <pj@sgi.com> 1.925.600.0401

_____

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