
Subject: Re: [PATCH 0/7] containers (V7): Generic Process Containers
Posted by [Paul Jackson](#) on Mon, 12 Feb 2007 09:18:43 GMT
[View Forum Message](#) <> [Reply to Message](#)

> - temporarily removed the "release agent" support.

ouch

> ... it can be re-added ... via a kernel thread that periodically polls containers ...

double ouch.

You'll have a rough time selling me on the idea that some kernel thread should be waking up every few seconds, grabbing system-wide locks, on a big honkin NUMA box, for the few times per hour, or less, that a cpuset is abandoned.

Offhand, that sounds mildly insane to me.

And how would this get the edge-triggered, rather than level-triggered, release? In other words, if a new cpuset is created, and marked with the notify_on_release flag, but otherwise not yet used (no child cpusets and no tasks in it) then it is not to be released (removed.) Only children and/or tasks are added, then later removed, is it a candidate for release. I guess you'll need yet another state bit, set when the cpuset is abandoned (last child removed or last pid exits/leaves), and cleared when this kernel thread visits the cpuset to see if it should be removed.

Can you explain to me how this intruded on the reference counting?

--

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