Subject: Re: [PATCH 0/7] containers (V7): Generic Process Containers
Posted by Paul Menage on Mon, 12 Feb 2007 09:32:51 GMT

View Forum Message <> Reply to Message

On 2/12/07, Paul Jackson <pj@sgi.com> wrote:

>
> 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.

I think it could be made smarter than that, e.g. have a workqueue task that's only woken when a refcount does actually reach zero. (I think that waking a workqueue task is something that can be done without too much worry about locks)

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

Essentially, it means that anything that releases a reference count on a container needs to be able to trigger a call to the release agent. The reference count is often released at a point when important locks are held, so you end up having to pass buffers into any function that might drop a ref count, in order to store a path to a release agent to be invoked.

In particular, the new container_clone() function can be called during the task fork path; at which point forking a new release_agent process would be impossible, or at least nasty. Additionally, if containers are potentially going to be used for virtual servers, having the release agent run from a top-level process rather than the process context that released the refcount sounds like a sane option.

Paul