Subject: Re: containers (was Re: -mm merge plans for 2.6.23)
Posted by Paul Menage on Tue, 10 Jul 2007 18:34:38 GMT
View Forum Message <> Reply to Message

On 7/10/07, Srivatsa Vaddagiri <vatsa@linux.vnet.ibm.com> wrote:
> >
> > Container stuff. Hold, I guess. I was expecting updates from Paul.
>
> Paul,
> Are you working on a new version? I thought it was mostly ready
> for mainline.
>

There are definitely some big changes that I want to make internally to the framework, but I guess they don't have to block pushing the basic framework to mainline.

I've got a new patchset that's primarily got all the various -mm fix patches rolled into the appropriate original patches, along with plus some small tweaks

- changed the Kconfig files to avoid using "select"
- adding the subsystem name as a prefix for each control file to enforce namespace scoping
- misc contributions from others

Short-term I also want to:

- rethink the linked list that runs through each task to its css_group object, since that seemed to hurt performance a bit, but for now that can probably be solved by just ripping it out and going back to scanning the tasklist to enumerate tasks in a container.

- extend the options parsing, so we can have more than just a list of subsystems. Probably changing the existing -o<subsys1>,<subsys2>,... to be one of:
  -osubsys=<subsys1>:<subsys2>:...,<otheropt>=<otherval>
  -osubsys=<subsys1>,subsys=<subsys2>,subsys=...,<otheropt>=<otherval>
  (what's the preferred convention for fs mount options with multiple values?)

I'd not realised that anything else depending on containers was ready for upstream merge, but if CFS group support is ready then merging a subset of them is probably a good idea, since this is an application that I can see a lot of people wanting to play with.

Andrew, how about we merge enough of the container framework to support CFS? Bits we could leave out for now include container_clone() support and the nsproxy subsystem, fork/exit callback hooks, and

possibly leave cpusets alone for now (which would also mean we could
skip the automatic release-agent stuff). I'm in Tokyo for the Linux
Foundation Japan symposium right now, but I should be able to get the
new patchset to you for Friday afternoon.

Paul

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

---

Subject: Re: containers (was Re: -mm merge plans for 2.6.23)
Posted by akpm on Tue, 10 Jul 2007 18:53:19 GMT
View Forum Message <> Reply to Message

On Tue, 10 Jul 2007 11:34:38 -0700
"Paul Menage" <menage@google.com> wrote:

> Andrew, how about we merge enough of the container framework to
> support CFS? Bits we could leave out for now include container_clone()
> support and the nsproxy subsystem, fork/exit callback hooks, and
> possibly leave cpusets alone for now (which would also mean we could
> skip the automatic release-agent stuff). I'm in Tokyo for the Linux
> Foundation Japan symposium right now, but I should be able to get the
> new patchset to you for Friday afternoon.

mm.. Given that you propose leaving bits out for the 2.6.23 merge, and
that changes are still pending and that nothing will _use_ the framework in
2.6.23 I'd be inclined to err on the side of caution and hold it all back
from 2.6.23.

This has the advantage that the merge will happen after the kernel-summit
containers discussion which I suspect will be an important point in the
life of this project...

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

---

Subject: Re: containers (was Re: -mm merge plans for 2.6.23)
Posted by Paul Menage on Tue, 10 Jul 2007 19:05:14 GMT
View Forum Message <> Reply to Message

On 7/10/07, Andrew Morton <akpm@linux-foundation.org> wrote:
> > Andrew, how about we merge enough of the container framework to

> > support CFS? Bits we could leave out for now include container_clone()
> > support and the nsproxy subsystem, fork/exit callback hooks, and
> > possibly leave cpusets alone for now (which would also mean we could
> > skip the automatic release-agent stuff). I'm in Tokyo for the Linux
> > Foundation Japan symposium right now, but I should be able to get the
> > new patchset to you for Friday afternoon.
>
> mm.. Given that you propose leaving bits out for the 2.6.23 merge, and
> that changes are still pending and that nothing will _use_ the framework in
> 2.6.23

That's what I was originally thinking too, but since CFS has been
merged, CFS group scheduling would use it.

Paul

Subject: Re: containers (was Re: -mm merge plans for 2.6.23)
Posted by Srivatsa Vaddagiri on Wed, 11 Jul 2007 04:55:16 GMT
View Forum Message <> Reply to Message

On Tue, Jul 10, 2007 at 11:53:19AM -0700, Andrew Morton wrote:
> On Tue, 10 Jul 2007 11:34:38 -0700
> "Paul Menage" <menage@google.com> wrote:
>
> > Andrew, how about we merge enough of the container framework to
> > support CFS? Bits we could leave out for now include container_clone()
> > support and the nsproxy subsystem, fork/exit callback hooks, and
> > possibly leave cpusets alone for now (which would also mean we could
> > skip the automatic release-agent stuff). I'm in Tokyo for the Linux
> > Foundation Japan symposium right now, but I should be able to get the
> > new patchset to you for Friday afternoon.
>
> mm.. Given that you propose leaving bits out for the 2.6.23 merge, and
> that changes are still pending and that nothing will _use_ the framework in
> 2.6.23 [...]

Andrew,
The cpu group scheduler is ready and waiting for the container patches
in 2.6.23 :)

Here are some options with us:

a. (As Paul says) merge enough of container patches to enable

its use with cfs group scheduler (and possibly cpusets?)

b. Enable group scheduling bits in 2.6.23 using the user-id grouping mechanism (aka fair user scheduler). For 2.6.24, we could remove this interface and use Paul's container patches instead. Since this means change of API interface between 2.6.23 and 2.6.24, I don't prefer this option.

c. Enable group scheduling bits only in -mm for now (2.6.23-mmX), using Paul's container patches. I can send you a short patch that hooks up cfs group scheduler with Paul's container infrastructure.

If a. is not possible, I would prefer c.

Let me know your thoughts ..

--
Regards,
vatsa

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