
Subject: Re: [RFC] [PATCH 0/3] containers: introduction

Posted by [serue](#) on Wed, 10 Jan 2007 21:42:46 GMT

[View Forum Message](#) <> [Reply to Message](#)

Quoting Serge E. Hallyn (serue@us.ibm.com):

> Following is a small patchset implementing what I was describing
> about earlier, namely semantics for a hierarchical container
> naming scheme.

>

> What works:

>

> 1. `ls -l /proc/$$/container`

> shows the full hierarchical name of the container;

>

> 2. `mount -t containerfs none /container`

> results in a file tree under `/container` representing the
> full container hierarchy

>

> 3. `cd /proc/$$/container; ls`

> results in a listing of child containers

>

> What doesn't work:

> 1. The `/proc/$$/container` link always appears dead (red
> in bash on my fedora test system) because it points
> into a `kern_mounted` fs.

>

> 2. Features like

>

> `cd /proc/$$/container`

> `mv container_3 my_child_container`

>

> to rename a container or

>

> `cd /proc/$$/container`

> `rm container_3`

>

> to kill all processes a container are unimplemented.

>

> 3. Semantics for entering a namespace are not only
> unimplemented, but entirely unconsidered thus far.

> I suppose one cool way to enter a container would
> be

>

> `ln -s /proc/$$/container/child_container /proc/$$/container`

>

> but that

>

> a. Does not provide the ability to switch only

- > some of the namespaces, as Herbert wants.
- > b. May be unimplementable using proc support
- > as is - not sure.

A conversation with Cedric today, we were thinking perhaps the way to achieve this is to create files under each container directory for each namespace type.

For instance,

```
d /containers/init_container/  
f /containers/init_container/network  
f /containers/init_container/uts  
f /containers/init_container/user  
f /containers/init_container/pid  
d /containers/init_container/vserver1/  
f /containers/init_container/vserver1/network  
f /containers/init_container/vserver1/uts  
f /containers/init_container/vserver1/user  
f /containers/init_container/vserver1/pid
```

Note that if I want to enter just the network namespace of vserver1, it's not quite right to say you're entering vserver1 at all, since it consists of each namespace therein. Rather, you might

```
mkdir /containers/init_container/vserver2  
ln -s /containers/init_container/vserver1/network \  
/containers/init_container/vserver2/  
echo /containers/init_container/vserver2 > /proc/$$/container  
exec /bin/sh
```

What happened? Well, we created a new container with no tasks. We linked vserver2's network namespace in there, then requested that we enter the container. Since no other namespaces had been linked in, all other namespaces will be inherited from our own namespace.

Thoughts?

-serge

Containers mailing list
Containers@lists.osdl.org
<https://lists.osdl.org/mailman/listinfo/containers>
