

---

Subject: containers development plans

Posted by [serge](#) on Mon, 02 Jul 2007 16:55:04 GMT

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

---

We are trying to create a roadmap for the next year of 'container' development, to be reported to the upcoming kernel summit. Containers here is a bit of an ambiguous term, so we are taking it to mean all of:

1. namespaces
2. process containers
3. checkpoint/restart

Naturally we can't actually predict what will and won't be worked on, let alone what will be going upstream. But the following is a list of features which it seems reasonable to think might be worked on next year:

1. completion of ongoing namespaces
  - pid namespace
  - net namespace
  - ro bind mounts
2. continuation with new namespaces
  - devpts, console, and ttydrivers
  - user
  - time
  - namespace management tools
  - namespace entering
3. any additional work needed for virtual servers?
  - i.e. in-kernel keyring usage for cross-namespace permissions, etc
4. task containers functionality
  - base features
  - specific containers
    - poll to see who has plans
5. checkpoint/restart
  - memory c/r
    - (there are a few designs and prototypes)
    - (though this may be ironed out by then)
    - per-container swapfile?
  - overall checkpoint strategy
    - in-kernel vs userspace-driven
  - overall restart strategy

What more needs to be added to this list?

A list of the people we are currently aware of who are showing interest in these features follows. What I'd like to know is, from this list, do some people know what general or specific areas they plan to or want to

work on over the next year?

Stakeholders:

Eric Biederman

google

ibm

kerlabs

openvz

osdl (Masahiko Takahashi?)

vserver

Who is missing from the list?

thanks,

-serge

---

Subject: Re: containers development plans

Posted by [Erich Focht](#) on Mon, 02 Jul 2007 19:15:06 GMT

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

---

Hi,

On Monday 02 July 2007 18:55, Serge E. Hallyn wrote:

> A list of the people we are currently aware of who are showing interest  
> in these features follows. What I'd like to know is, from this list, do  
> some people know what general or specific areas they plan to or want to  
> work on over the next year?

>

> Stakeholders:

> Eric Biederman

> google

> ibm

> kerlabs

> openvz

> osdl (Masahiko Takahashi?)

> vserver

>

> Who is missing from the list?

you could add the XtreamOS project to the list. It's grid project and we need to use containers for grid job isolation. We're currently using BLCR for checkpointing, and also kerrighed, but aim at switching to containers and will work on container checkpointing. XtreamOS is related to kerlabs & kerrighed, but has broader interests in containers.

> What more needs to be added to this list?

By the way, there are potential synergies between container checkpointing work and the need of the VM to have something like per cpuset of per container swap (as explained in the VM BoFS at OLS, there are unnecessary OOMs with a common swap).

Regards,  
Erich

---

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

---

---

Subject: Re: containers development plans  
Posted by [serue](#) on Mon, 02 Jul 2007 21:53:15 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Quoting Erich Focht (efocht@hpce.nec.com):

> Hi,  
>  
> On Monday 02 July 2007 18:55, Serge E. Hallyn wrote:  
> > A list of the people we are currently aware of who are showing interest  
> > in these features follows. What I'd like to know is, from this list, do  
> > some people know what general or specific areas they plan to or want to  
> > work on over the next year?  
> >  
> > Stakeholders:  
> > Eric Biederman  
> > google  
> > ibm  
> > kerlabs  
> > openvz  
> > osdl (Masahiko Takahashi?)  
> > vserver  
> >  
> > Who is missing from the list?  
>  
> you could add the XtreamOS project to the list. It's grid project and we  
> need to use containers for grid job isolation. We're currently using BLCR  
> for checkpointing, and also kerrighed, but aim at switching to containers  
> and will work on container checkpointing. XtreamOS is related to  
> kerlabs & kerrighed, but has broader interests in containers.

That's cool. I'll go read up on XtreamOS before I ask stupid questions...

> > What more needs to be added to this list?

>  
> By the way, there are potential synergies between container checkpointing  
> work and the need of the VM to have something like per cpuset or per  
> container swap (as explained in the VM BoFS at OLS, there are unnecessary  
> OOMs with a common swap).

Ah, cool, I hadn't heard that. At the risk of really exploding the # of container subsystems needed to be composed in order to do checkpoint/restart, this does sound like a separate container subsystem which just attaches a swapfile to it's tasks might then be a useful thing. It can then be composed with a cpuset or a freezer subsystem, or both (at the same time - separately might be weird :).

I guess then for checkpoint we just freeze the subsystem, force a write of all dirty pages using something like Dave's patch, somehow create a copy-on-write clone as the checkpoint, then unfreeze. (and oh yeah, checkpoint all the other crap too :)

> Regards,  
> Erich

thanks,  
-serge

---

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

---

---

Subject: netns summary (was Re: containers development plans)  
Posted by [Cedric Le Goater](#) on Thu, 05 Jul 2007 14:19:08 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Serge E. Hallyn wrote:

> We are trying to create a roadmap for the next year of  
> 'container' development, to be reported to the upcoming kernel  
> summit. Containers here is a bit of an ambiguous term, so we are  
> taking it to mean all of:  
>  
> 1. namespaces  
> 2. process containers  
> 3. checkpoint/restart  
>  
> Naturally we can't actually predict what will and won't be worked on,  
> let alone what will be going upstream. But the following is a list  
> of features which it seems reasonable to think might be worked on  
> next year:  
>

- > 1. completion of ongoing namespaces
- > pid namespace
- > net namespace

I'm not sure if this has been said already :

At OLS, we had a talk with denis, pavel, eric, daniel, benjamin and others, and we agreed to cooperate on eric's netns patchset which seems to satisfy most of the stake holders :

- \* openvz
- \* ibm
- \* eric (hopefully :)
- \* planetlab (heavy users of linux-vserver)
- \* hp (also interested)
- \* google (?)

the planetlab team successfully included eric's netns patchset in the linux-vserver patch and had good results with it.

the following tasks were discussed :

- \* improve the patchset to make it acceptable by the netdev community
- \* share a netns git tree
- \* share some tests framework all parties have been developing independently.

Are we in sync ? I'm sure there are more interesting stuff to be said :)

We're looking for a place to host the netns git tree while we are working on it, which means we will need first a federator for the patchset. Eric declined as he is too busy. Any proposals ?

Thanks,

C.

---

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

---

---

Subject: Re: containers development plans  
Posted by [Cedric Le Goater](#) on Thu, 05 Jul 2007 15:53:37 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Hello,

some more comments on what we talked about at OLS and what we are working on.

Serge E. Hallyn wrote:

- > We are trying to create a roadmap for the next year of
- > 'container' development, to be reported to the upcoming kernel
- > summit. Containers here is a bit of an ambiguous term, so we are
- > taking it to mean all of:
- >
- > 1. namespaces
- > 2. process containers
- > 3. checkpoint/restart
- >
- > Naturally we can't actually predict what will and won't be worked on,
- > let alone what will be going upstream. But the following is a list
- > of features which it seems reasonable to think might be worked on
- > next year:
- >
- > 1. completion of ongoing namespace

the ipc namespace would need a "set identifier" feature if we were to use it for C/R. this is not available right now. a patchset was sent introducing a new IPC\_SETID but it didn't get much attention.

- > pid namespace

At OLS, we agreed that suka's hierarchical pidns patchset should be fine if we can make sure perms are OK when the namespace is not used. right ?

I get < 1% today, so it should be okay :)

There are still some issues around /proc that we are working on. Hopefully, we should be able to merge most of the helpers patch real soon.

we need a clone\_with\_pid() kind of syscall for C/R. I had planned to work on a :

```
clone64(struct clone64_arg_struct *arg)
```

to extend the clone flags which will soon overflow. we could easily add a pid attribute to implement the clone\_with\_pid()

the kthread cleanup is not completed yet. some patch are pending but i would say that the most important ones are around NFS and i'm not sure anyone worked on these.

af unix credentials still hold some pid\_t's. they need a clean up.

> net namespace

see previous email

> ro bind mounts

work in progress. dave ?

what about mounting /proc and /sys multiple times ?

> 2. continuation with new namespaces

> devpts, console, and ttydrivers

> user

merged experimental. we still need to work on the (user,users) checks. however, openvz and linux-vserver should already be able to use it.

> time

> namespace management tools

> namespace entering

there are a few patchsets on the topic :

\* bind\_ns() syscall

\* container subsystem identifying a nsproxy object

but they didn't get much review :(

> 3. any additional work needed for virtual servers?

> i.e. in-kernel keyring usage for cross-username namespace permissions, etc

> 4. task containers functionality

> base features

> specific containers

> poll to see who has plans

> 5. checkpoint/restart

we really need to leverage the freezer and suspend to disk for that. there are some talks about it right now but it seems a bit early to have clear directions yet.

generalizing the refrigerator to all arch seems a good idea to freeze a container. then how do we initiate checkpoint ? syscall ? signal ? etc.

These topics were addressed at the BOF and people are now aware of different solutions. we hope that the email storm on what directions

to take for mainline will start soon.

C.

- > memory c/r
- > (there are a few designs and prototypes)
- > (though this may be ironed out by then)
- > per-container swapfile?
- > overall checkpoint strategy
- > in-kernel vs userspace-driven

hybrid ?

- > overall restart strategy
- >
- > What more needs to be added to this list?
- >
- > A list of the people we are currently aware of who are showing interest
- > in these features follows. What I'd like to know is, from this list, do
- > some people know what general or specific areas they plan to or want to
- > work on over the next year?
- >
- > Stakeholders:
- > Eric Biederman
- > google
- > ibm
- > kerlabs
- > openvz
- > osdl (Masahiko Takahashi?)
- > vserver
- >
- > Who is missing from the list?
- >
- > thanks,
- > -serge
- >
- > \_\_\_\_\_
- > Containers mailing list
- > Containers@lists.linux-foundation.org
- > <https://lists.linux-foundation.org/mailman/listinfo/containers>
- >

---

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

---



Subject: Re: containers development plans  
Posted by [Paul Jackson](#) on Thu, 05 Jul 2007 15:56:22 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Could someone explain how these directions impact Paul Menage's container patches and, the one I care about the most, kernel/cpuset.c?

--

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

---

Containers mailing list  
[Containers@lists.linux-foundation.org](mailto:Containers@lists.linux-foundation.org)  
<https://lists.linux-foundation.org/mailman/listinfo/containers>

---

---

Subject: Re: containers development plans  
Posted by [Cedric Le Goater](#) on Thu, 05 Jul 2007 16:00:29 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Paul Jackson wrote:

> Could someone explain how these directions impact Paul Menage's  
> container patches and, the one I care about the most, kernel/cpuset.c?

they are indirectly related. Sorry for the noise.

Paul Menage's container patches provide a process aggregation mechanism, like PAGG did. This is a common requirement for resource management and other container features like resource isolation and checkpoint/restart.

Cheers,

C.

---

Containers mailing list  
[Containers@lists.linux-foundation.org](mailto:Containers@lists.linux-foundation.org)  
<https://lists.linux-foundation.org/mailman/listinfo/containers>

---

---

Subject: Re: containers development plans  
Posted by [Paul Jackson](#) on Thu, 05 Jul 2007 16:06:34 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

> they are indirectly related. Sorry for the noise.

Well ... if this is noise, it should be me apologizing, not you <grin>.

I'm the one who asked.

> Paul Menage's container patches provide a process aggregation mechanism,  
> like PAGG did. This is a common requirement for resource management and  
> other container features like resource isolation and checkpoint/restart.

So ... will you (those using the netns patches) be using Paul Menage's containers, or competing with them for the same ends, or what?

To repeat my original question, what is the relation between the work described on this thread (which some OLS discussions have decided to base on the netns patches) and Paul Menage's containers and my (now the community's) cpusets.

--

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>

---

---

Subject: Re: containers development plans  
Posted by [ebiederm](#) on Thu, 05 Jul 2007 17:23:48 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Paul Jackson <pj@sgi.com> writes:

>> they are indirectly related. Sorry for the noise.  
>  
> Well ... if this is noise, it should be me apologizing, not you <grin>.  
> I'm the one who asked.  
>  
>> Paul Menage's container patches provide a process aggregation mechanism,  
>> like PAGG did. This is a common requirement for resource management and  
>> other container features like resource isolation and checkpoint/restart.  
>  
> So ... will you (those using the netns patches) be using Paul Menage's  
> containers, or competing with them for the same ends, or what?  
>  
> To repeat my original question, what is the relation between the work  
> described on this thread (which some OLS discussions have decided to  
> base on the netns patches) and Paul Menage's containers and my (now the  
> community's) cpusets.

Largely orthogonal.

This discussions is hard because Paul has appropriated our term for the user space aggregation of all of the pieces (a container) using for some subset of that, so my apologies if there is some confusion.

There are a few issues with the filesystem part of Paul's patchset that currently do not allow for nested containers.

Generally the Paul's container filesystem work provides a generic framework for resources controls that we want in addition to the namespaces. I'm not at all certain I like the filesystem interface to user space, but having a common interface to user space and helper code to use it for all of the resource controls makes sense.

I think there is a little more like a user space visible identifier that is also interesting.

Eric

---

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

---

---

Subject: Re: containers development plans  
Posted by [Paul Jackson](#) on Thu, 05 Jul 2007 17:41:57 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Thanks for responding, Eric.

For those like me (I suspect I'm not alone) who drift in and out of this discussion, perhaps this would be a good time to describe again what are the containers you're discussing here - their key purpose(s) and their essential architecture.

For extra credit, putting aside the conflicting use of the word 'container', how would, in your view, PMC's (Paul Menage's Containers) relate to TTC's (This Thread's Containers)? Would either use the other? Are they complimentary, conflicting, ...?

--

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

---

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

---

Subject: Re: containers development plans

Posted by [Herbert Poetzl](#) on Mon, 09 Jul 2007 12:12:57 GMT

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

---

On Mon, Jul 02, 2007 at 11:55:04AM -0500, Serge E. Hallyn wrote:

> We are trying to create a roadmap for the next year of  
> 'container' development, to be reported to the upcoming kernel  
> summit. Containers here is a bit of an ambiguous term, so we are  
> taking it to mean all of:

- >
- > 1. namespaces
- > 2. process containers
- > 3. checkpoint/restart

>

> Naturally we can't actually predict what will and won't be worked on,  
> let alone what will be going upstream. But the following is a list  
> of features which it seems reasonable to think might be worked on  
> next year:

- >
- > 1. completion of ongoing namespaces
  - > pid namespace
  - > net namespace
  - > ro bind mounts
- > 2. continuation with new namespaces
  - > devpts, console, and ttydrivers
  - > user
  - > time

various accounting and limits including disk space  
and memory

- > namespace management tools
- > namespace entering

well, that is currently covered by the Linux-VServer  
interfaces (and I assume by the OpenVZ API too), but  
I guess you are more interested in a mainline tool  
and interface

- > 3. any additional work needed for virtual servers?
  - > i.e. in-kernel keyring usage for cross-username namespace permissions, etc

nfs and rpc, as well as general security/auth  
frameworks have to be adjusted ...

- > 4. task containers functionality
  - > base features
  - > specific containers

- > poll to see who has plans
- > 5. checkpoint/restart
- > memory c/r
- > (there are a few designs and prototypes)
- > (though this may be ironed out by then)
- > per-container swapfile?
- > overall checkpoint strategy
- > in-kernel vs userspace-driven
- > overall restart strategy
- >
- > What more needs to be added to this list?
- >
- > A list of the people we are currently aware of who are showing interest
- > in these features follows. What I'd like to know is, from this list, do
- > some people know what general or specific areas they plan to or want to
- > work on over the next year?
- >
- > Stakeholders:
- > Eric Biederman
- > google
- > ibm
- > kerlabs
- > openvz
- > osdl (Masahiko Takahashi?)
- > vserver
- >
- > Who is missing from the list?

Linux-VServer as usual :)

best,  
Herbert

- > thanks,
- > -serge

---

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

---

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

---

Subject: Re: netns summary (was Re: containers development plans)  
Posted by [Herbert Poetzl](#) on Mon, 09 Jul 2007 12:16:14 GMT

On Thu, Jul 05, 2007 at 04:19:08PM +0200, Cedric Le Goater wrote:

> Serge E. Hallyn wrote:

> > We are trying to create a roadmap for the next year of  
> > 'container' development, to be reported to the upcoming kernel  
> > summit. Containers here is a bit of an ambiguous term, so we are  
> > taking it to mean all of:

> >

> > 1. namespaces

> > 2. process containers

> > 3. checkpoint/restart

> >

> > Naturally we can't actually predict what will and won't be worked on,  
> > let alone what will be going upstream. But the following is a list  
> > of features which it seems reasonable to think might be worked on  
> > next year:

> >

> > 1. completion of ongoing namespaces

> > pid namespace

> > net namespace

>

> I'm not sure if this has been said already :

>

> At OLS, we had a talk with denis, pavel, eric, daniel, benjamin and  
> others, and we agreed to cooperate on eric's netns patchset which  
> seems to satisfy most of the stake holders :

> \* openvz

> \* ibm

> \* eric (hopefully :)

> \* planetlab (heavy users of linux-vserver)

> \* hp (also interested)

> \* google (?)

>

> the planetlab team successfully included eric's netns patchset in the  
> linux-vserver patch and had good results with it.

except for the increased overhead (just for the record)  
(so functionality wise, it seemed fine)

> the following tasks were discussed :

>

> \* improve the patchset to make it acceptable by the netdev community

> \* share a netns git tree

> \* share some tests framework all parties have been developing  
> independently.

>

> Are we in sync ? I'm sure there are more interesting stuff to be said :)

>  
> We're looking for a place to host the netns git tree while we are  
> working on it, which means we will need first a federator for the  
> pachset. Eric declined as he is too busy. Any proposals ?

interesting ...

best,  
Herbert

> Thanks,

>

> C.

>

---

> Containers mailing list

> Containers@lists.linux-foundation.org

> <https://lists.linux-foundation.org/mailman/listinfo/containers>

---

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>

---

Subject: Re: containers development plans  
Posted by [serge](#) on Mon, 09 Jul 2007 15:17:22 GMT  
[View Forum Message](#) <> [Reply to Message](#)

Quoting Paul Jackson (pj@sgi.com):

> > they are indirectly related. Sorry for the noise.

>

> Well ... if this is noise, it should be me apologizing, not you <grin>.

> I'm the one who asked.

>

> > Paul Menage's container patches provide a process aggregation mechanism,

> > like PAGG did. This is a common requirement for resource management and

> > other container features like resource isolation and checkpoint/restart.

>

> So ... will you (those using the netns patches) be using Paul Menage's

> containers, or competing with them for the same ends, or what?

>

> To repeat my original question, what is the relation between the work

> described on this thread (which some OLS discussions have decided to

Hi Paul,

containers does mean two things in this thread, but note that we  
want to discuss both meanings.

So if any development is planned over the next year on top of Pauls Menage's containers and/or cpusets, we'd like that listed here as well.

thanks,  
-serge

> base on the netns patches) and Paul Menage's containers and my (now the  
> community's) cpusets.

>

> --

> 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>

---

---

Subject: Re: containers development plans  
Posted by [Paul Jackson](#) on Mon, 09 Jul 2007 15:23:20 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Serge wrote:

> So if any development is planned over the next year on top of  
> Pauls Menage's containers and/or cpusets, we'd like that listed  
> here as well.

Well, apparently you guys are planning some "container" stuff that is on top of or relates to in some way unclear to me with Menage's containers and/or cpusets.

I'll repeat my question a third time ... not knowing any clearer way to state it than before, I'll just quote myself:

> So ... will you (those using the netns patches) be using Paul Menage's  
> containers, or competing with them for the same ends, or what?

>

> To repeat my original question, what is the relation between the work  
> described on this thread (which some OLS discussions have decided to  
> base on the netns patches) and Paul Menage's containers and my (now the  
> community's) cpusets.

--

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



---

Subject: Re: containers development plans  
Posted by [serge](#) on Mon, 09 Jul 2007 15:25:58 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Quoting Paul Jackson (pj@sgi.com):

> Thanks for responding, Eric.

>

> For those like me (I suspect I'm not alone) who drift in and out of  
> this discussion, perhaps this would be a good time to describe again  
> what are the containers you're discussing here - their key purpose(s)  
> and their essential architecture.

>

> For extra credit, putting aside the conflicting use of the word  
> 'container', how would, in your view, PMC's (Paul Menage's Containers)  
> relate to TTC's (This Thread's Containers)? Would either use the other?

Again, we want to discuss both in this roadmap.

Let's stick to talking about PMCs and namespaces (TTC's).

The namespace work just aims to take any resources which userspace might wish to identify by a global id, and allow a process to unshare it's view of that namespace so that

1. it can isolate its resources from those of other processes (virtual servers)
2. it can guarantee that certain ids are available, to make process restart possible.

PMCs are very useful to both process restart and virtual servers:

1. for resource management
2. for providing APIs for namespace operations
  - a. for example, the freezer subsystem Cedric recently sent out, which can be used to freeze a checkpoint/restart job or virtual server (or any type of container) to allow safe checkpoint.
  - b. to restrict a virtual server's cpu and memory usage

> Are they complimentary, conflicting, ...?

Very much complimentary.

thanks,  
-serge

---

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

---

---

Subject: Re: containers development plans  
Posted by [serge](#) on Mon, 09 Jul 2007 15:27:40 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Quoting Paul Jackson (pj@sgi.com):

> Serge wrote:  
> > So if any development is planned over the next year on top of  
> > Pauls Menage's containers and/or cpusets, we'd like that listed  
> > here as well.  
>  
> Well, apparenntly you guys are planning some "container" stuff  
> that is on top of or relates to in some way unclear to me with  
> Menage's containers and/or cpusets.  
>  
> I'll repeat my question a third time ... not knowing any clearer  
> way to state it than before, I'll just quote myself:

Please see the next response I sent, and let me know what I've still failed to clarify there.

(trying to avoid more repeating :)

thanks,  
-serge

> > So ... will you (those using the netns patches) be using Paul Menage's  
> > containers, or competing with them for the same ends, or what?  
> >  
> > To repeat my original question, what is the relation between the work  
> > described on this thread (which some OLS discussions have decided to  
> > base on the netns patches) and Paul Menage's containers and my (now the  
> > community's) cpusets.  
>  
> --  
> I won't rest till it's the best ...  
> Programmer, Linux Scalability  
> Paul Jackson <pj@sgi.com> 1.925.600.0401

---

Containers mailing list

---

Subject: Re: [ckrm-tech] containers development plans  
Posted by [dev](#) on Mon, 09 Jul 2007 15:32:58 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Paul Jackson wrote:

> Serge wrote:

>

>>So if any development is planned over the next year on top of  
>>Pauls Menage's containers and/or cpusets, we'd like that listed  
>>here as well.

>

>

> Well, apparently you guys are planning some "container" stuff  
> that is on top of or relates to in some way unclear to me with  
> Menage's containers and/or cpusets.

>

> I'll repeat my question a third time ... not knowing any clearer  
> way to state it than before, I'll just quote myself:

>

>

>>So ... will you (those using the netns patches) be using Paul Menage's  
>>containers, or competing with them for the same ends, or what?

>>

>>To repeat my original question, what is the relation between the work  
>>described on this thread (which some OLS discussions have decided to  
>>base on the netns patches) and Paul Menage's containers and my (now the  
>>community's) cpusets.

Paul,

Paul Menage's containers is a good interfaces for grouping tasks  
which is planned to be used for resource controls (including cpusets which actually  
control cpu resources as well). So it is planned to be used for:  
cpusets, RSS, fair CPU and other resource controls.

Network namespaces, ipc and other namespaces use unshare() mechanisms  
and can have a simple view in containers fs.

Thanks,  
Kirill

---

Containers mailing list  
Containers@lists.linux-foundation.org

---

Subject: Re: containers development plans  
Posted by [Paul Menage](#) on Tue, 10 Jul 2007 05:56:56 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

On 7/2/07, Serge E. Hallyn <[serge@hallyn.com](mailto:serge@hallyn.com)> wrote:

> 4. task containers functionality

How about if we adopt "process containers" or "task containers" as the term for the generic container framework, to distinguish from more general user-space containers? In the same way that "task\_struct" in the kernel is understood to be separate from the concept of a "task" in a job scheduling system in userspace.

> base features

Features that I'd like to see in the short and medum term:

- support for virtualized containerfs mounts, so that virtual servers can mount their own containerfs and manage sub-containers
- automatically prefixing control file names with the subsystem name, unless changed or disabled by the user at mount time
- removing unnecessary locking where possible.
- simplifying the control file API
- a userspace RBCE along with simple configuration so that you can easily use generic containers to apply subsystem controls on a per-user, per-group, per-pgrp, per-executable, etc, basis. (E.g. to easily apply CFS to be fair between pgrps rather than fair between processes)

> specific containers

> poll to see who has plans

Some possible subsystems that I'm thinking of include:

- splitting the memory and cpu isolation parts of cpusets into two separate subsystems (still backwards-compatible)
- some kind of network connect/bind/accept controller. Eric came up with a nice way of doing this by adding iptables hooks for connect/bind/accept, and then adding an iptables match module that could match based on container id. This would give us all the

flexibility of iptables and the existing iptables tools. The drawback is that it could be rather tricky to virtualize. A less flexible solution that just allowed you to specify permitted local-port-range/remote-port-range/remote-netmask tuples would be more virtualizable, even if it doesn't make as much reuse of existing iptables support.

- some way of controlling which network flow ids (used as inputs into standard Linux queueing) processes in a container can use.
- userspace per-container OOM handler, maybe as part of cpusets or some other memory controller.

Paul

---

---

Subject: Re: [ckrm-tech] containers development plans  
Posted by [Balbir Singh](#) on Tue, 10 Jul 2007 06:01:22 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Paul Menage wrote:  
[snip]

- > Some possible subsystems that I'm thinking of include:
- >
- > - splitting the memory and cpu isolation parts of cpusets into two
- > separate subsystems (still backwards-compatible)

I see memory isolation using cpusets as very topology dependent and I am not sure if the model would work for memory controllers.

- > - userspace per-container OOM handler, maybe as part of cpusets or
- > some other memory controller.
- >

Yes, I think it would be nice to get a userspace OOM handler in (into the main system first) and then adapt it for containers.

- > Paul
- >

--

Warm Regards,  
Balbir Singh  
Linux Technology Center  
IBM, ISTL

---

---

Subject: Re: [ckrm-tech] containers development plans  
Posted by [Paul Menage](#) on Tue, 10 Jul 2007 06:19:05 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

On 7/9/07, Balbir Singh <balbir@linux.vnet.ibm.com> wrote:  
> > - splitting the memory and cpu isolation parts of cpusets into two  
> > separate subsystems (still backwards-compatible)  
>  
> I see memory isolation using cpusets as very topology dependent  
> and I am not sure if the model would work for memory controllers.

I wasn't suggesting making any changes to the page-based memory controllers as part of this.

Currently in the mainline kernel, the cpumask and nodemask portions of cpusets are essentially two mostly-independent modules that happen to be coupled together in the same file and use the same process tracking system (cpusets). Once we have generic process containers, splitting this into a "cpusets" subsystem that handles all the cpumask portions of the existing cpusets, and a "memsets" subsystem that handles all the nodemask and memory migration portions would remove that coupling and give more flexibility.

Paul

---

---

Subject: Re: [ckrm-tech] containers development plans  
Posted by [Balbir Singh](#) on Tue, 10 Jul 2007 06:59:16 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Paul Menage wrote:  
> On 7/9/07, Balbir Singh <balbir@linux.vnet.ibm.com> wrote:  
>> > - splitting the memory and cpu isolation parts of cpusets into two  
>> > separate subsystems (still backwards-compatible)  
>>  
>> I see memory isolation using cpusets as very topology dependent  
>> and I am not sure if the model would work for memory controllers.  
>  
> I wasn't suggesting making any changes to the page-based memory  
> controllers as part of this.  
>  
> Currently in the mainline kernel, the cpumask and nodemask portions of  
> cpusets are essentially two mostly-independent modules that happen to  
> be coupled together in the same file and use the same process tracking  
> system (cpusets). Once we have generic process containers, splitting  
> this into a "cpusets" subsystem that handles all the cpumask portions  
> of the existing cpusets, and a "memsets" subsystem that handles all  
> the nodemask and memory migration portions would remove that coupling

> and give more flexibility.

>

> Paul

Aaah.. I see, that makes sense from a cpusets/containers perspective.

--

Warm Regards,  
Balbir Singh  
Linux Technology Center  
IBM, ISTL

---

---

Subject: Re: [ckrm-tech] containers development plans  
Posted by [Paul Jackson](#) on Tue, 10 Jul 2007 07:25:20 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Kirill, Serge, et al,

Is it fair to say then that Paul Menage's containers are primarily for the purposes of managing resources, while namespaces are for the purposes of managing identifiers?

We've got some resources, like cpu cycles, memory bytes, network bandwidth, that we want to allocate and account for differentially by groups of tasks -- that's Menage's containers.

We've got some system wide namespaces, like process id's, that we want to virtualize, for more flexible uses -- these are the namespace containers.

In Serge's opening post to this thread, he wrote:

1. namespaces
2. process containers
3. checkpoint/restart

Are the 'process containers' of item (2) the containers of Paul Menage?

If so, then I propose that this thread is misnamed. It should not be "containers development plans", but rather "namespace, container and c/r development plans." And if so, there is really no conflict over the use of the word 'container' -- that applies just to the resource virtualization efforts, of which my cpusets is the granddaddy example, being generalized by Paul Menage with his container patches. The other work is, as Serge actually termed it in the body of his post, better called 'namespaces'.

Perhaps the confusion arose from looking for a single word to encompass all three parts, listed above, of this work. The efforts have some strong dependencies, but taking the name of one of the efforts, containers, and trying to make it serve double duty as the umbrella term, might be an unnecessary confusion.

Perhaps also, on a separate point, the word 'process' in item (2) is not the right focus. I think that the essential purpose of (2) is resource management. While most of those resource management is done per-process, it might also be per-file or per-virtual-address-range. For example, disk i/o backing pages in a particular range of an applications address space might have certain bandwidth limits, and the memory backing the pages in that range might have certain memory node placement restrictions, or the i/o to and from a particular disk file might have certain bandwidth or placement constraints. We see a bit of this in present day kernels, with the mbind(2) system call.

So, I suggest, we have three efforts:

1. namespaces,
2. resource containers, and
3. checkpoint/restart.

And they are very much in this order. Names, such as pathnames, task pids, user uids, and other system wide identifiers, are essential to the implementation of all else. Resource containers depend on some naming scheme, and reach out to manage the use of resources outside the operating system, such as disk, network, memory and processor. Checkpoint/restart is a particular feature of interest, that requires that both names and resources be virtualized to some degree.

Are there any Python programmers in the namespace work? The use of namespaces in Python might serve as a informative example for the work we need in Linux namespaces. See further page 418, section "A.2 Namespaces and Binds" of David Mertz's "Text Processing in Python" for a clear and concise exposition of the central role of namespaces in Python.

--

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>

---

---



Subject: Re: [ckrm-tech] containers development plans  
Posted by [Paul Jackson](#) on Tue, 10 Jul 2007 07:26:54 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Balbir wrote:

> Aaah.. I see, that makes sense from a cpuset/containers perspective.

Agreed.

--

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

---

---

Subject: Re: containers development plans  
Posted by [Paul Jackson](#) on Tue, 10 Jul 2007 07:38:08 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Paul M wrote:

> How about if we adopt "process containers" or "task containers" as the  
> term for the generic container framework, to distinguish from more  
> general user-space containers?

As I just spent the last hour writing in another reply (not noticing your message in my queue), I suspect that these are not best called 'process' (or 'task') containers, but rather 'resource' containers.

You provide what look like more good examples of why this is so, with your mention of various proposals for managing network traffic, which (I'm unsure of this) seem not to bind per-task, but per-link or some such.

Do the various 'subsystems' that you're thinking of correspond to the various resources that we're virtualizing?

Separate question -- what do you mean by 'the generic container framework' ... I'm clueless.

--

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

---

---

Subject: Re: [ckrm-tech] containers development plans  
Posted by [Paul Menage](#) on Tue, 10 Jul 2007 07:53:20 GMT

---

On 7/10/07, Paul Jackson <pj@sgi.com> wrote:

> Kirill, Serge, et al,

>

> Is it fair to say then that Paul Menage's containers are primarily  
> for the purposes of managing resources, while namespaces are for the  
> purposes of managing identifiers?

Sort of - but one thing that we're trying to figure out how to do nicely is integrate namespaces into the container filesystem (this was the purpose of the post\_clone() container API callback) so that we can both get a filesystem view of task namespaces, and combine namespaces with other process container subsystems.

>

> We've got some resources, like cpu cycles, memory bytes, network  
> bandwidth, that we want to allocate and account for differentially  
> by groups of tasks -- that's Menage's containers.

Plus things that aren't necessarily resource controllers, such as the container freezer, or permissions on network ports, or userspace OOM handlers. I don't think that lumping all of these in together as "resource containers" is the right thing to do.

> virtualization efforts, of which my cpusets is the granddaddy example,  
> being generalized by Paul Menage with his container patches. The other  
> work is, as Serge actually termed it in the body of his post, better  
> called 'namespaces'.

Purely within the kernel, yes. The more general encompassing effort to have a combined kernel/userspace solution for virtual servers is also referred to as "containers". (And to be fair that term was already in use when I started using the term "process containers" to refer to the specific framework in the kernel that handles process tracking).

Paul

---

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>

---

---

Subject: Re: [ckrm-tech] containers development plans  
Posted by [Paul Jackson](#) on Tue, 10 Jul 2007 08:30:37 GMT

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

---

Paul M wrote:

> Sort of - but ...

Thanks for your reply.

I guess I'll just have to wander off, still a tad confused.

I suspect my instincts are running in a contrary direction to those, including yourself, doing the work here. My preference is to pick a subset of the task before me that I can clearly "brand", with a name and description that others can quickly understand. Better to clearly express 80% of the goal in a well branded product, than to confuse the 'customers' (those you want to buy into to the proposal) with a conceptually more elaborate, but less well defined, attempt to cover all 100% of the goal.

Often times, the quickest way to convey a complex idea is to convey an overly simplified approximation, and then allow the listener to tell you why what you said is too simple minded. This is because the bandwidth within a brain vastly exceeds that between two brains.

But since I'm not volunteering to do the work, I should not waste too much more time of those who are.

--

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>

---

---

Subject: Re: [ckrm-tech] containers development plans  
Posted by [serge](#) on Tue, 10 Jul 2007 16:32:12 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Quoting Paul Jackson (pj@sgi.com):

> Kirill, Serge, et al,

>

> Is it fair to say then that Paul Menage's containers are primarily  
> for the purposes of managing resources, while namespaces are for the  
> purposes of managing identifiers?

>

> We've got some resources, like cpu cycles, memory bytes, network  
> bandwidth, that we want to allocate and account for differentially  
> by groups of tasks -- that's Menage's containers.

>

> We've got some system wide namespaces, like process id's, that we  
> want to virtualize, for more flexible uses -- these are the name-  
> space containers.  
>  
> In Serge's opening post to this thread, he wrote:  
> 1. namespaces  
> 2. process containers  
> 3. checkpoint/restart  
>  
> Are the 'process containers' of item (2) the containers of Paul Menage?

Yup.

> If so, then I propose that this thread is misnamed. It should not be  
> "containers development plans", but rather "namespace, container and  
> c/r development plans." And if so, there is really no conflict over  
> the use of the word 'container' -- that applies just to the resource  
> virtualization efforts, of which my cpusets is the granddaddy example,  
> being generalized by Paul Menage with his container patches. The other  
> work is, as Serge actually termed it in the body of his post, better  
> called 'namespaces'.  
>  
> Perhaps the confusion arose from looking for a single word to encompass  
> all three parts, listed above, of this work. The efforts have some strong

Not exactly - the "confusion" arose because the ksummit committee wanted to hear about "containers", and agreed that by that term they mean each of those three. So I kept the term 'containers' in the roadmap title, but we can change that if it's preferred.

thanks,  
-serge

---

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

---

Subject: Re: containers development plans  
Posted by [serge](#) on Tue, 10 Jul 2007 20:33:26 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Quoting Herbert Poetzl (herbert@13thfloor.at):  
> On Mon, Jul 02, 2007 at 11:55:04AM -0500, Serge E. Hallyn wrote:  
> > We are trying to create a roadmap for the next year of  
> > 'container' development, to be reported to the upcoming kernel  
> > summit. Containers here is a bit of an ambiguous term, so we are  
> > taking it to mean all of:

> >  
> > 1. namespaces  
> > 2. process containers  
> > 3. checkpoint/restart  
> >  
> > Naturally we can't actually predict what will and won't be worked on,  
> > let alone what will be going upstream. But the following is a list  
> > of features which it seems reasonable to think might be worked on  
> > next year:  
> >  
> > 1. completion of ongoing namespaces  
> > pid namespace  
> > net namespace  
> > ro bind mounts  
>  
> > 2. continuation with new namespaces  
> > devpts, console, and ttydrivers  
> > user  
> > time  
>  
> various accounting and limits including disk space  
> and memory

Is it fair to leave that under task container subsystems?

> > namespace management tools  
> > namespace entering  
>  
> well, that is currently covered by the Linux-VServer  
> interfaces (and I assume by the OpenVZ API too), but  
> I guess you are more interested in a mainline tool  
> and interface

I'm interested in anything someone is willing to try to push upstream :)

> > 3. any additional work needed for virtual servers?  
> > i.e. in-kernel keyring usage for cross-username namespace permissions, etc  
>  
> nfs and rpc, as well as general security/auth  
> frameworks have to be adjusted ...  
>  
> > 4. task containers functionality  
> > base features  
> > specific containers  
> > poll to see who has plans  
> > 5. checkpoint/restart  
> > memory c/r

> > (there are a few designs and prototypes)  
> > (though this may be ironed out by then)  
> > per-container swapfile?  
> > overall checkpoint strategy  
> > in-kernel vs userspace-driven  
> > overall restart strategy  
> >  
> > What more needs to be added to this list?  
> >  
> > A list of the people we are currently aware of who are showing interest  
> > in these features follows. What I'd like to know is, from this list, do  
> > some people know what general or specific areas they plan to or want to  
> > work on over the next year?  
> >  
> > Stakeholders:  
> > Eric Biederman  
> > google  
> > ibm  
> > kerlabs  
> > openvz  
> > osdl (Masahiko Takahashi?)  
> > vserver  
> >  
> > Who is missing from the list?  
>  
> Linux-VServer as usual :)

I stand corrected as usual.

-serge

---

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

---

---

Subject: Re: [ckrm-tech] containers development plans  
Posted by [serge](#) on Tue, 10 Jul 2007 21:30:45 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Quoting Serge E. Hallyn ([serge@hallyn.com](mailto:serge@hallyn.com)):  
> Quoting Paul Jackson ([pj@sgi.com](mailto:pj@sgi.com)):  
> > Kirill, Serge, et al,  
> >  
> > Is it fair to say then that Paul Menage's containers are primarily  
> > for the purposes of managing resources, while namespaces are for the  
> > purposes of managing identifiers?  
> >

> > We've got some resources, like cpu cycles, memory bytes, network  
> > bandwidth, that we want to allocate and account for differentially  
> > by groups of tasks -- that's Menage's containers.  
> >  
> > We've got some system wide namespaces, like process id's, that we  
> > want to virtualize, for more flexible uses -- these are the name-  
> > space containers.  
> >  
> > In Serge's opening post to this thread, he wrote:  
> > 1. namespaces  
> > 2. process containers  
> > 3. checkpoint/restart  
> >  
> > Are the 'process containers' of item (2) the containers of Paul Menage?  
>  
> Yup.  
>  
> > If so, then I propose that this thread is misnamed. It should not be  
> > "containers development plans", but rather "namespace, container and  
> > c/r development plans." And if so, there is really no conflict over  
> > the use of the word 'container' -- that applies just to the resource  
> > virtualization efforts, of which my cpusets is the granddaddy example,  
> > being generalized by Paul Menage with his container patches. The other  
> > work is, as Serge actually termed it in the body of his post, better  
> > called 'namespaces'.  
> >  
> > Perhaps the confusion arose from looking for a single word to encompass  
> > all three parts, listed above, of this work. The efforts have some strong  
>  
> Not exactly - the "confusion" arose because the ksummit committee wanted  
> to hear about "containers", and agreed that by that term they mean each  
> of those three. So I kept the term 'containers' in the roadmap title,  
> but we can change that if it's preferred.

I plan to keep the thread titled 'containers' precisely because \*I\*  
don't care whose work gets renamed, while several other people on both  
sides care so strongly, so it would seem rude for me to make that  
decision de-facto in this way.

Maybe renaming one or both projects should be listed in the roadmap as a  
todo :)

thanks,  
-serge

---

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

---

Subject: Re: containers development plans  
Posted by [dev](#) on Thu, 12 Jul 2007 10:32:27 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Paul Menage wrote:

> On 7/2/07, Serge E. Hallyn <[serge@hallyn.com](mailto:serge@hallyn.com)> wrote:

>

>> 4. task containers functionality

>

>

> How about if we adopt "process containers" or "task containers" as the  
> term for the generic container framework, to distinguish from more  
> general user-space containers? In the same way that "task\_struct" in  
> the kernel is understood to be separate from the concept of a "task"  
> in a job scheduling system in userspace.

>

>

>> base features

>

>

> Features that I'd like to see in the short and medum term:

>

> - support for virtualized containerfs mounts, so that virtual servers  
> can mount their own containerfs and manage sub-containers

>

> - automatically prefixing control file names with the subsystem name,  
> unless changed or disabled by the user at mount time

>

> - removing unnecessary locking where possible.

>

> - simplifying the control file API

>

> - a userspace RBCE along with simple configuration so that you can  
> easily use generic containers to apply subsystem controls on a  
> per-user, per-group, per-pgrp, per-executable, etc, basis. (E.g. to  
> easily apply CFS to be fair between pgrps rather than fair between  
> processes)

>

>

>> specific containers

>> poll to see who has plans

>

>

> Some possible subsystems that I'm thinking of include:

>

> - splitting the memory and cpu isolation parts of cpuset into two  
> separate subsystems (still backwards-compatible)

>

> - some kind of network connect/bind/accept controller. Eric came up



> with a nice way of doing this by adding iptables hooks for  
> connect/bind/accept, and then adding an iptables match module that  
> could match based on container id. This would give us all the  
> flexibility of iptables and the existing iptables tools. The drawback  
> is that it could be rather tricky to virtualize. A less flexible  
> solution that just allowed you to specify permitted  
> local-port-range/remote-port-range/remote-netmask tuples would be more  
> virtualizable, even if it doesn't make as much reuse of existing  
> iptables support.

Not sure why it requires some additional controller, but surely  
it is possible to create a match for iptables matching container ID.  
Is it what you are thinkinh about or I got something wrong?

Thanks,  
Kirill

---

---

Subject: Re: containers development plans  
Posted by [Paul Menage](#) on Thu, 12 Jul 2007 10:44:43 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

On 7/12/07, Kirill Korotaev <dev@sw.ru> wrote:

>  
> Not sure why it requires some additional controller, but surely  
> it is possible to create a match for iptables matching container ID.

But which container ID? Don't forget that a task is in one container  
in each hierarchy of which there could be more than one. At its  
simplest this new subsystem could just be a way to tell iptables which  
hierarchy to look at when matching based on container id. In practice  
it's probably reasonable to make the "iptables container id"  
user-settable since userspace is building the iptables rules and might  
want to use its own numbering scheme for the ids. (E.g. all container  
IDs in a particular range have the same kinds of permissions).

Paul

---

---

Subject: Re: [ckrm-tech] containers development plans  
Posted by [dev](#) on Thu, 12 Jul 2007 10:58:36 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Paul Menage wrote:

> On 7/12/07, Kirill Korotaev <dev@sw.ru> wrote:  
>  
>>Not sure why it requires some additional controller, but surely

>>it is possible to create a match for iptables matching container ID.  
>  
>  
> But which container ID? Don't forget that a task is in one container  
> in each hierarchy of which there could be more than one. At its  
> simplest this new subsystem could just be a way to tell iptables which  
> hierarchy to look at when matching based on container id. In practice  
> it's probably reasonable to make the "iptables container id"  
> user-settable since userspace is building the iptables rules and might  
> want to use its own numbering scheme for the ids. (E.g. all container  
> IDs in a particular range have the same kinds of permissions).

won't hierarchy:container-name pair help? :@)

Kirill

---

---

Subject: Re: containers development plans  
Posted by [kir](#) on Sat, 14 Jul 2007 03:21:05 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

I just got an idea -- what about organizing a one-day "containers mini-summit" just before the Kernel Summit? We can all meet face to face, discuss all the issues and come out with a good plan for KS talk. That does not mean we shouldn't discuss it here, of course -- it's just to make it "polished".

I guess organizers of LinuxConf Europe (which takes place just before the KS) will have a facility (i.e. a room) for us.

What do you think?

Serge E. Hallyn wrote:

> We are trying to create a roadmap for the next year of  
> 'container' development, to be reported to the upcoming kernel  
> summit.

---

---

Subject: Re: containers development plans  
Posted by [serge](#) on Sat, 14 Jul 2007 20:15:37 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Quoting Kir Kolyshkin (kir@openvz.org):

> I just got an idea -- what about organizing a one-day "containers  
> mini-summit" just before the Kernel Summit? We can all meet face to  
> face, discuss all the issues and come out with a good plan for KS talk.  
> That does not mean we shouldn't discuss it here, of course -- it's just

> to make it "polished".  
>  
> I guess organizers of LinuxConf Europe (which takes place just before  
> the KS) will have a facility (i.e. a room) for us.  
>  
> What do you think?

I like the idea, but I'm not sure it's practical since I suspect most people who aren't going to kernel summit won't be able to get funding to go to the mini-summit. I could be wrong about that...

How many people will be at LinuxConf Europe - can we see a show of hands?

thanks,  
-serge

> Serge E. Hallyn wrote:  
> >We are trying to create a roadmap for the next year of  
> >'container' development, to be reported to the upcoming kernel  
> >summit.  
>

---

Subject: Re: containers development plans  
Posted by [kir](#) on Sat, 14 Jul 2007 23:28:03 GMT  
[View Forum Message](#) <> [Reply to Message](#)

Serge E. Hallyn wrote:  
> Quoting Kir Kolyshkin (kir@openvz.org):  
>  
>> I just got an idea -- what about organizing a one-day "containers  
>> mini-summit" just before the Kernel Summit? We can all meet face to  
>> face, discuss all the issues and come out with a good plan for KS talk.  
>> That does not mean we shouldn't discuss it here, of course -- it's just  
>> to make it "polished".  
>>  
>> I guess organizers of LinuxConf Europe (which takes place just before  
>> the KS) will have a facility (i.e. a room) for us.  
>>  
>> What do you think?  
>>  
>  
> I like the idea, but I'm not sure it's practical since I suspect most  
> people who aren't going to kernel summit won't be able to get funding to  
> go to the mini-summit. I could be wrong about that...  
>  
> How many people will be at LinuxConf Europe - can we see a show of

> hands?

>

Me and Pavel Emelyanov -- since we both have talks on the LinuxConf.

---

---

Subject: Re: containers development plans

Posted by [Rohit Seth](#) on Mon, 16 Jul 2007 23:08:31 GMT

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

---

On Sat, 2007-07-14 at 05:21 +0200, Kir Kolyshkin wrote:

> I just got an idea -- what about organizing a one-day "containers  
> mini-summit" just before the Kernel Summit? We can all meet face to  
> face, discuss all the issues and come out with a good plan for KS talk.

Do you know who is doing the presentation in KS for containers?

>

> That does not mean we shouldn't discuss it here, of course -- it's just  
> to make it "polished".

>

> I guess organizers of LinuxConf Europe (which takes place just before  
> the KS) will have a facility (i.e. a room) for us.

>

> What do you think?

I think this is a good idea. There will be couple folks from Google if this is arranged.

Cheers,  
-rohit

---

---

Subject: Re: containers development plans

Posted by [serge](#) on Tue, 17 Jul 2007 10:17:12 GMT

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

---

Quoting Rohit Seth (rohitseth@google.com):

> On Sat, 2007-07-14 at 05:21 +0200, Kir Kolyshkin wrote:

> > I just got an idea -- what about organizing a one-day "containers  
> > mini-summit" just before the Kernel Summit? We can all meet face to  
> > face, discuss all the issues and come out with a good plan for KS talk.

>

> Do you know who is doing the presentation in KS for containers?

Nope, that is not decided yet, but we should know soon.

> > That does not mean we shouldn't discuss it here, of course -- it's just

> > to make it "polished".  
> >  
> > I guess organizers of LinuxConf Europe (which takes place just before  
> > the KS) will have a facility (i.e. a room) for us.  
> >  
> > What do you think?  
>  
> I think this is a good idea. There will be couple folks from Google if  
> this is arranged.

Unfortunately I'm unlikely to be there (late notice - I really need to keep a better list of upcoming conferences and proposal deadlines!). But it still sounds like a good idea.

Depending on the format of the containers mini-summit, would it make sense to set up a phone number so others can dial in? Or would it be better to have a smaller group anyway?

-serge

---

Subject: Re: containers development plans  
Posted by [dev](#) on Tue, 17 Jul 2007 13:51:37 GMT  
[View Forum Message](#) <> [Reply to Message](#)

> Unfortunately I'm unlikely to be there (late notice - I really need to  
> keep a better list of upcoming conferences and proposal deadlines!).  
> But it still sounds like a good idea.

It's a pity :/ I though you are the primary candidate for presenting containers at KLS.

> Depending on the format of the containers mini-summit, would it make  
> sense to set up a phone number so others can dial in? Or would it be  
> better to have a smaller group anyway?

I think we can do both :) after having a mini-summit to discuss some pre-summary over the phone.

Regards,  
Kirill

---

Subject: Re: [ckrm-tech] containers development plans  
Posted by [Paul Menage](#) on Tue, 17 Jul 2007 16:19:22 GMT  
[View Forum Message](#) <> [Reply to Message](#)

On 7/17/07, Serge E. Hallyn <[serge@hallyn.com](mailto:serge@hallyn.com)> wrote:

>  
> Depending on the format of the containers mini-summit, would it make  
> sense to set up a phone number so others can dial in? Or would it be  
> better to have a smaller group anyway?  
>

I'd be interested in calling in if I can't make it myself.

Paul

---

Subject: Re: containers development plans  
Posted by [serue](#) on Tue, 17 Jul 2007 18:10:02 GMT  
[View Forum Message](#) <> [Reply to Message](#)

Quoting Kirill Korotaev (dev@sw.ru):

> > Unfortunately I'm unlikely to be there (late notice - I really need to  
> > keep a better list of upcoming conferences and proposal deadlines!).  
> > But it still sounds like a good idea.  
>  
> It's a pity :/ I though you are the primary candidate for presenting containers at KLS.

I would be able to get approval to attend KS. I just wouldn't be able to get approval to attend the linuxconf europe without presenting a paper, which it is too late to do.

So if I happen to go to KS, I can attend the miniconf. If not, then:

> > Depending on the format of the containers mini-summit, would it make  
> > sense to set up a phone number so others can dial in? Or would it be  
> > better to have a smaller group anyway?  
>  
> I think we can do both :) after having a mini-summit to discuss some pre-summary  
> over the phone.

Sounds great. (I'd hoped to do this for the OLS bof, but that didn't work out :( )

thanks,  
-serge

---

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

---

Subject: Re: [ckrm-tech] containers development plans

Posted by [serue](#) on Tue, 17 Jul 2007 18:11:49 GMT

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

---

Quoting Paul (?\$BJuN) Menage (menage@google.com):

> On 7/17/07, Serge E. Hallyn <serge@hallyn.com> wrote:

> >

> > Depending on the format of the containers mini-summit, would it make

> > sense to set up a phone number so others can dial in? Or would it be

> > better to have a smaller group anyway?

> >

>

> I'd be interested in calling in if I can't make it myself.

Great - if someone brings an IP phone with which they can dial toll free US numbers, I have a conference # we can use.

-serge

---

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>

---

---

Subject: Re: containers development plans

Posted by [Cedric Le Goater](#) on Mon, 23 Jul 2007 13:40:06 GMT

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

---

Serge E. Hallyn wrote:

> Quoting Kirill Korotaev (dev@sw.ru):

>>> Unfortunately I'm unlikely to be there (late notice - I really need to

>>> keep a better list of upcoming conferences and proposal deadlines!).

>>> But it still sounds like a good idea.

>> It's a pity :/ I though you are the primary candidate for presenting containers at KLS.

>

> I would be able to get approval to attend KS. I just wouldn't be able to

> get approval to attend the linuxconf europe without presenting a paper,

> which it is too late to do.

>

> So if I happen to go to KS, I can attend the miniconf. If not, then:

when is the mini conf planned ? the 3rd of September, just before KS ?

thanks,

C.

---

Containers mailing list

Containers@lists.linux-foundation.org

---

Subject: Re: containers development plans  
Posted by [serge](#) on Mon, 23 Jul 2007 14:10:00 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Quoting Cedric Le Goater (clg@fr.ibm.com):

> Serge E. Hallyn wrote:

> > Quoting Kirill Korotaev (dev@sw.ru):

> >>> Unfortunately I'm unlikely to be there (late notice - I really need to  
> >>> keep a better list of upcoming conferences and proposal deadlines!).

> >>> But it still sounds like a good idea.

> >> It's a pity :/ I though you are the primary candidate for presenting containers at KLS.

> >

> > I would be able to get approval to attend KS. I just wouldn't be able to  
> > get approval to attend the linuxconf europe without presenting a paper,  
> > which it is too late to do.

> >

> > So if I happen to go to KS, I can attend the miniconf. If not, then:

>

> when is the mini conf planned ? the 3rd of September, just before KS ?

That would definately be my hope if i were going to ks

-serge

---

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

---