
Subject: Re: [RFC][PATCH 2.6.22-rc5] System V IPC: new IPC_SETID command to modify an ID

Posted by [serue](#) on Wed, 20 Jun 2007 20:19:04 GMT

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

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

> Kirill Korotaev wrote:

> > Cedric Le Goater wrote:

> >> Pierre Peiffer wrote:

> >>

> >>

> >>> This patch adds a new IPC_SETID command to the System V IPCs set of
> >>> commands, which allows to change the ID of an existing IPC.

> >>>

> >>> This command can be used through the semctl/shmctl/msgctl API, with the new
> >>> ID passed as the third argument for msgctl and shmctl (instead of a
> >>> pointer) and through the fourth argument for semctl.

> >>>

> >>> To be successful, the following rules must be respected:

> >>> - the IPC exists

> >>> - the user must be allowed to change the IPC attributes regarding the IPC
> >>> permissions.

> >>> - the new ID must satisfy the ID computation rule.

> >>> - the entry (in the kernel internal table of IPCs) corresponding to the new
> >>> ID must be free.

> >>

> >> That's an interesting way to reset the ids of sysv ipcs during a restart (after
> >> a checkpoint) and we're looking for ways to do that among other things.

> >>

> >> How does it fit openvz ? Is it something openvz could use ?

> >

> > my personal imho is that we should not export such interfaces to user space
> > and do the checkpointing from the kernel.

> > it simplifies a lot of things and makes checkpointing more elegant.

You say more elegant, someone else might say less flexible and more
kernel version dependant.

I guess we can discuss this more at the bof next thursday, but this
question, whether to have the checkpoint be kernel-guided, or userspace
guided, is one we really need to talk about.

> I agree that we should not export useless interface to user but even

If we use it to guide checkpointing from the kernel, then none of it is
useless :)

> without exporting anything we need support from the kernel internals

> (sysvipc in that case) to allow checkpointing. like setting the id.
> right ?
>
> so there might be some common ground with some new user space interface.
> That's what we are looking for.

Pierre, do you want to break out the internal support (either from your patchset, or from openvz's) and post that separately? Then the small patch on top of that to either present the set_id functionality to userspace or to use it in a kernel-guided restore should be trivial.

thanks,
-serge

> > So until there is some user-space usage scenario of the patch I missed -
> > i wouldn't commit it.
>
> Pierre has a user space tool to use it. quite simple.
>
> 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>
