

Hi,

A couple of months ago, Pierre Peiffer has submitted a patch series to enable checkpointing / restarting ipc.

4 of these patches were related to semaphores:

<https://lists.linux-foundation.org/pipermail/containers/2008-January/thread.html#9756>  
<https://lists.linux-foundation.org/pipermail/containers/2008-January/thread.html#9757>  
<https://lists.linux-foundation.org/pipermail/containers/2008-January/thread.html#9758>  
<https://lists.linux-foundation.org/pipermail/containers/2008-January/thread.html#9759>

They introduced a new procfs file: `/proc/<pid>/semundo` to read and write the semaphores undo values for a given process.

These patches are widely used in the -lxc development tree cryo code is based upon.

(more information about lxc can be found at <http://lxc.sourceforge.net/> and the development tree can be found there too - under the patches link).

Manfred Spraul, on his side, has rewritten an important part of the semaphores code. See:

1. <http://lkml.org/lkml/2008/5/24/92>
2. <http://lkml.org/lkml/2008/5/24/93>
3. <http://lkml.org/lkml/2008/5/24/90>
4. <http://lkml.org/lkml/2008/5/24/91>

Mainly patches 1 and 4 made Pierre's patches unappliable.

- . patch 1 changed the semundo\_list proc\_list field into a linked list.
- . patch 2
  - . reversed the locking order of the sem\_undo\_list lock and the semaphore lock.
  - . converted the sem\_undo structure to use rcu.

Since 2.6.26-rc5-mm3 is now the new target for the -lxc development tree, I've ported them and taking this opportunity to resubmit them to the containers list.

I have in mind a simpler solution, which I think I'll propose next week: instead of writing into the `/proc/<pid>/semundo` of a third party process, only allow the write operation to be done for `<current>`.

Kathy, can you please pull in these patches?

Regards,  
Nadia

--

---

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

---