

---

Subject: [RFC PATCH 0/6] SYSVIPC/semaphores - allow saving/restoring a process' semundo\_list

Posted by [Nadia Derby](#) on Wed, 25 Jun 2008 13:49:10 GMT

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

---

This patchset is a part of an effort to make sysv ipc objects read/writable from userspace for checkpoint / restart.

System V ipc's are objects that are global to a system and can thus be checkpointed and restarted on a container basis. But some parts of the ipc structures are process related and, as such should be checkpointed / restarted on a process basis.

Message queues needn't and thus cannot be reached from a task structure (only the other direction is possible: a task can be reached by a msg receiver or sender structure if it is sleeping).

Shared memories are accessible from a task structure through that task's memory mapping (/proc/<pid>/maps shows a process' memory maps).

Semaphores are kind of accessible from a task structure too: the task structure's sysvsem field makes it possible to walk through all the semaphores operations to undo when a process is exiting.

This list, that need to be saved and restored during a process' c/r, cannot yet be accessed from user space.

This is a feature that will be needed if ever we take the direction of driving checkpoint / restart from user space, though the read part of it could be used even from now on.

Since this undo\_list is, again, on a thread basis, we propose to externalize it via a new proc file: /proc/<pid>/semundo.

Actually, Pierre Pieffer has already done the proposal in threads <https://lists.linux-foundation.org/pipermail/containers/2008-January/thread.html#9756> up to #9759

I've ported them to 2.6.26-rc5-mm3, and I'm now coming back with a simpler implementation: the write operation is now only allowed into /proc/self/semundo, which simplifies the locking strategy.

These patches should be applied to 2.6.26-rc5-mm3, in the following order:

[ PATCH 1/6 ] : ipc\_rcu\_protect\_access\_to\_undo\_list.patch

Makes the process' undo\_list rcu protected in order to enable safely reading it.

[ PATCH 2/6 ] : ipc\_procfs\_semundo\_file.patch

Introduces the semundo proc file (the seq operations are still

empty).

[ PATCH 3/6] : ipc\_procfs\_semundo\_start\_stop\_seqops.patch

Introduces the .start and .stop seq operations.

[ PATCH 4/6] : ipc\_procfs\_semundo\_next\_seqop.patch

Introduces the .next seq operation.

[ PATCH 5/6] : ipc\_procfs\_semundo\_show\_seqop.patch

Introduces the .show seq operation.

[ PATCH 6/6] : ipc\_procfs\_semundo\_write.patch

The semundo proc file becomes writable.

Comments are welcome!

Regards,

Nadia

--

---

Containers mailing list

Containers@lists.linux-foundation.org

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

---