
Subject: [PATCH 0/20] Pid namespaces (release candidate)
Posted by [Pavel Emelianov](#) on Tue, 07 Aug 2007 09:28:17 GMT
[View Forum Message](#) <> [Reply to Message](#)

This is the version that is ready to be sent to Andrew. Oleg advised not to do it right now, so I'm sending it to containers@ for final review and (mainly) severe testing. If anyone wants one may add its Acked/Signed-off marks to patches :)

Changelog from the previous version:

- * fixed exit_task_namespaces() with Oleg's patch
- * fixed init exiting
- * removed the work_struct from pid namespace
- * added kmem cache for pid namespaces
- * fixed namespaces termination
- * optimized and fixed namespaces reference counting
- * split this set for git-bisect
- * split this set into smaller pieces
- * removed signal handling

A pid namespace is a "view" of a particular set of tasks on the system. They work in a similar way to filesystem namespaces. A file (or a process) can be accessed in multiple namespaces, but it may have a different name in each. In a filesystem, this name might be /etc/passwd in one namespace, but /chroot/etc/passwd in another.

For processes, a process may have pid 1234 in one namespace, but be pid 1 in another. This allows new pid namespaces to have basically arbitrary pids, and not have to worry about what pids exist in other namespaces. This is essential for checkpoint/restart where a restarted process's pid might collide with an existing process on the system's pid.

In this particular implementation, pid namespaces have a parent-child relationship, just like processes. A process in a pid namespace may see all of the processes in the same namespace, as well as all of the processes in all of the namespaces which are children of its namespace. Processes may not, however, see others which are in their parent's namespace, but not in their own. The same goes for sibling namespaces.

This set is based on my patches, I sent before, but it includes some comments and patches that I received from Sukadev. Sukadev, please, add your Acked-by, Signed-off-by or From, to patches you want (everybody is also welcome :)).

Signed-off-by: Pavel Emelianov <xemul@openvz.org>

Signed-off-by: Sukadev Bhattiprolu <sukadev@us.ibm.com>
