Subject: [RFC] Virtualization patches for IPC/UTS. 2nd step
Posted by dev on Fri, 24 Mar 2006 17:23:39 GMT
View Forum Message <> Reply to Message

I propose to consider these patches for utsname and sysv IPC
virtualization once more.

The main change from the previous version is an introduction of separate
namespaces for each subsytem as suggested by Eric Biederman. People who
work on containers don't care actually, so I splitted a container into
namespaces.

The naming conventions used in these patches are (as example for IPC):

CONFIG_IPC_NS - per namespace config option
struct ipc_namespace - structure describing namepsace
ipc_ns   - names of var pointers, in task_struct etc.
init_ipc_ns - default host namespace

interfaces:
get_ipc_ns - refcountig interface
put_ipc_ns - refcountig interface
create_ipc_ns - create _empty_ namespace
clone_ipc_ns - clone current namespace, if applicable
free_ipc_ns - destroy namespace when refs are 0

Please, note, these patches do not introduce CONFIG_XXX_NS option in any
of Kconfigs, as it is questionable whether to have them scattered all
around or place in some menu "Virtual namespaces". But patches compile
and work fine both w/o and with it.

Also, please, note, that these patches do not virtualize corresponding
sysctls or proc parts of uts/ipc. I suppose this must be postponed as we
have no any consensus on /proc.

Some other minor differences from Eric/Dave patches:
- these patches heavily use current namespace context instead of
  bypassing additional argument to all functions where required.
- these patches compile to the old good kernel when namespaces are off.

Both patches are also available in GIT repo at:
http://git.openvz.org/pub/linux-2.6-openvz-ms/

Thanks,
Kirill