Subject: Re: [PATCH 1/9] namespaces: add nsproxy
Posted by serue on Mon, 22 May 2006 12:39:19 GMT
View Forum Message <> Reply to Message

Quoting Eric W. Biederman (ebiederm@xmission.com):
> Sam Vilain <sam@vilain.net> writes:
>
> > Serge E. Hallyn wrote:
> >
> >>@@ -1585,7 +1591,15 @@ asmlinkage long sys_unshare(unsigned lon
> >>
> >> if (new_fs || new_ns || new_sigh || new_mm || new_fd || new_ulist) {
> >>
> >>+ old_nsproxy = current->nsproxy;
> >>+ new_nsproxy = dup_namespaces(old_nsproxy);
> >>+ if (!new_nsproxy) {
> >>+   err = -ENOMEM;
> >>+   goto bad_unshare_cleanup_semundo;
> >>+ }
> >>+
> >>   task_lock(current);
> >>
> >>
> >
> > We'll get lots of duplicate nsproxy structures before we move all of the
> > pointers for those subsystems into it. Do we need to dup namespaces on
> > all of those conditions?
>
> Ugh. Good catch. The new nsproxy needs to be just for the fs and the uts
> namespace.
>
> I guess that means that test should be moved up a few lines.

Oh. Yeah. It didn't look odd to me bc it's about the number of
namespaces we are *going* to have :)

Fix follows:

Subject: [PATCH] uts: copy nsproxy only when needed.
From: Serge Hallyn <serue@us.ibm.com>

The nsproxy was being copied in unshare() when anything was being
unshared, even if it was something not referenced from nsproxy.
This should end up in some cases with far more memory usage than
necessary.

Signed-off-by: Serge Hallyn <serue@us.ibm.com>

```
---

 kernel/fork.c |   20 ++++++++++++++------
 1 files changed, 14 insertions(+), 6 deletions(-)

74d1068458c62302ac8ed38e38a57b692580662f
diff --git a/kernel/fork.c b/kernel/fork.c
index cdc549e..9278a68 100644
--- a/kernel/fork.c
+++ b/kernel/fork.c
@@ -1559,7 +1559,7 @@ asmlinkage long sys_unshare(unsigned lon
  struct mm_struct *mm, *new_mm = NULL, *active_mm = NULL;
  struct files_struct *fd, *new_fd = NULL;
  struct sem_undo_list *new_ulist = NULL;
- struct nsproxy *new_nsproxy, *old_nsproxy;
+ struct nsproxy *new_nsproxy = NULL, *old_nsproxy = NULL;
  struct uts_namespace *uts, *new_uts = NULL;

  check_unshare_flags(&unshare_flags);
@@ -1587,18 +1587,24 @@ asmlinkage long sys_unshare(unsigned lon
  if ((err = unshare_utsname(unshare_flags, &new_uts)))
   goto bad_unshare_cleanup_semundo;

- if (new_fs || new_ns || new_sigh || new_mm || new_fd || new_ulist ||
-     new_uts) {
-
+ if (new_ns || new_uts) {
   old_nsproxy = current->nsproxy;
   new_nsproxy = dup_namespaces(old_nsproxy);
   if (!new_nsproxy) {
    err = -ENOMEM;
    goto bad_unshare_cleanup_uts;
   }
+ }
+
+ if (new_fs || new_ns || new_sigh || new_mm || new_fd || new_ulist ||
+     new_uts) {

   task_lock(current);
-  current->nsproxy = new_nsproxy;
+
+  if (new_nsproxy) {
+   current->nsproxy = new_nsproxy;
+   new_nsproxy = old_nsproxy;
+  }

   if (new_fs) {
    fs = current->fs;
```

```
@@ -1640,9 +1646,11 @@ asmlinkage long sys_unshare(unsigned lon
   }

   task_unlock(current);
-  put_nsproxy(old_nsproxy);
  }

+ if (new_nsproxy)
+  put_nsproxy(new_nsproxy);
+
 bad_unshare_cleanup_uts:
  if (new_uts)
   put_uts_ns(new_uts);
--
1.1.6
```