Subject: Re: [PATCH] utsns: Restore proper namespace handling.
Posted by serue on Tue, 27 Nov 2007 16:28:32 GMT

Quoting Andrew Morton (akpm@linux-foundation.org):
> On Mon, 26 Nov 2007 09:19:17 -0600 "Serge E. Hallyn" <serue@us.ibm.com> wrote:
>
> > Quoting Eric W. Biederman (ebiederm@xmission.com):
> > >
> > > When CONFIG_UTS_NS was removed it seems that we also deleted
> > > the code for handling sysctls in the other then the initial
> > > uts namespace.   This patch restores that code.
> > >
> > > Signed-off-by: Eric W. Biederman <ebiederm@xmission.com>
> >
> > Thanks, Eric.
> >
> > Acked-by: Serge Hallyn <serue@us.ibm.com>
> >
> > > ---
> > > kernel/utsname_sysctl.c | 2 ++
> > > 1 files changed, 2 insertions(+), 0 deletions(-)
> > >
> > > diff --git a/kernel/utsname_sysctl.c b/kernel/utsname_sysctl.c
> > > index c76c064..71f58c3 100644
> > > --- a/kernel/utsname_sysctl.c
> > > +++ b/kernel/utsname_sysctl.c
> > > @@ -18,6 +18,8 @@
> > > static void *get_uts(ctl_table *table, int write)
> > > {
> > >   char *which = table->data;
> > > + struct uts_namespace *uts_ns = current->nsproxy->uts_ns;
> > > + which = (which - (char *)&init_uts_ns) + (char *)uts_ns;
> > >
> > >   if (!write)
> > >   down_read(&uts_sem);
>
> I already have a (more codingstylely attractive) version of this from
> Pavel, for which I shall steal your ack.

Cool, thanks.

> --- a/kernel/utsname_sysctl.c~isolate-the-uts-namespaces-domainname-and-hostname-back
> +++ a/kernel/utsname_sysctl.c
> @@ -18,6 +18,10 @@
> static void *get_uts(ctl_table *table, int write)
> {
>   char *which = table->data;

> + struct uts_namespace *uts_ns;
> +
> + uts_ns = current->nsproxy->uts_ns;
> + which = (which - (char *)&init_uts_ns) + (char *)uts_ns;
>
>   if (!write)
>   down_read(&uts_sem);
> _
>
>
> Those pointer tricksies are revolting. What's going on in there?

This is true. There is only one sysctl table, and it's data points
to the location of the desired field in the initial namespace. So
the trickery finds the same field in the target namespace.

I believe Eric and Pavel are working on a patchset that will make
this trickery unnecessary by having sort of shadow sysctls.

-serge

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