Subject: Re: cryo and mm->arg_start
Posted by Matt Helsley on Fri, 11 Jul 2008 22:01:13 GMT
View Forum Message <> Reply to Message

On Fri, 2008-07-11 at 09:38 -0700, Dave Hansen wrote:
> On Fri, 2008-07-11 at 08:13 -0500, Serge E. Hallyn wrote:
> >
> > One thing we could do here is to start extending the cryo approach
> > with Eric's checkpoint-as-a-coredump (caac?). We generate the
> > tiniest of coredumps which, at first, contains nothing but
> > mm->arg_start and maybe a process id. It would be simplest if
> > it also contained a filename for the real executable,
>
> The exec model sounds reasonable to me.
>
> But, I think the filename of the exe is going to have to be in the
> checkpoint *already*. It is mapped by at least one of the VMAs, and
> will probably be dumped as a normal file-backed area.

Yes, the file that backed the exec will be there. Note that thanks to "stacking" filesystems the path to the file backing the exe is not _always_ going to be the same as the path to the file which userspace exec'd in the first place. You can see this by comparing the /proc/<pid>/exe symlink with the file backing the VMA.

This is important to any program which checks the /proc/self/exe symlink to find out where it's installed (Java does this, for example). I think it's possible to do this with a binfmt -- it's just one more detail to remember.

Cheers,
-Matt