

"Serge E. Hallyn" <serue@us.ibm.com> writes:

> How about we solve both this and the general ugliness of using ptrace
> with a new
>
> hijack_and_clone(struct task_struct *tsk, int fd)
>
> Which takes tsk, clones it, and execs the contents of fd?

That is what roughly what I was thinking. Although that is an ugly beast to implement. Getting stdin and stdout should still be doable using a tty. Getting the semantics and the implementation right is a tough challenge.

After thinking about it a normal exec (unless you want your binary to come from inside the namespace) is nearly useless because it requires a static binary. With glibc not actually going static, static binaries are nearly impossible to write. Although that might be a good argument for minimalism, and security.

The really important use of the ptrace case is that it works using existing mechanisms without leaks. So it is very useful yardstick.

The other important yardstick is arranging it so that when you login to a machine all of the user code runs in your target environment. How you get there is irrelevant.

One of the cases I have been worrying about in looking at the semantics of enter is what do you do with the parent pid. Supporting ptrace from outside the pid namespace of a process inside a pid namespace requires supporting a parent process outside of the pid namespace for processes other than init.

I'm not convinced setting up a non-ptrace parent that is outside the pid namespace makes sense, but it looks like the mechanism is going to be there. If we did support a foreign parent it would go a long way towards supporting the login and be redirected into a container case, as well as Herbert's pid namespace without an init case.

I have finally worked through all of the reasonable irq handling alternatives and unless something goes wrong will I will be submitting that code tomorrow. I really want to pull some pid namespace patches together so we can bring those into the conversation but I don't think I will be able to get there before I head out first of March

to Nebraska to spend some time with my brother.

Eric

Containers mailing list

Containers@lists.osdl.org

<https://lists.osdl.org/mailman/listinfo/containers>
