
Subject: Re: [RFC PATCH 0/5] Resend -v2 - Use procfs to change a syscall behavior

Posted by [ebiederm](#) on Thu, 10 Jul 2008 02:14:37 GMT

[View Forum Message](#) <> [Reply to Message](#)

Alexey Dobriyan <adobriyan@gmail.com> writes:

> In case of ->did_exec the only clean interface I see is:
>
> tsk->did_exec = !tsk_img->did_exec;
>
> It would be pretty silly to wrap this one line in a system call (two
> actually -- one in, one out), since you're going to restore some more
> fields of such variety anyway (like ->pdeath_signal).

There I agree the granularity seems small enough to be a major pain for the implementation.

> Given the diversity of kernel internal data structures and all sorts of
> links between them, the only system call suitable is ioctl(2), not all
> this zoo of system calls proposed. They are all extendable and without
> rules, but ioctl(2) is also without rules.

At least for processes my gut reaction is to look at binary formats and coredumps. Something with at least that large of a granularity seems to make most sense.

> This is all said in assumption that serializing kernel-internal data for
> checkpoint/restart to userspace is acceptable for mainline.
> I don't think it is.

I don't believe that serializing kernel-internal data is acceptable for mainline. I believe that serializing user-visible data is acceptable.
Note: user-visible data does not mean user-manipulatable data.

On a socket you may not save the skbs but you can save the pending packets for example. Assuming the transition cost is not too high.

Eric

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