
Subject: Re: Extending syscalls

Posted by [hpa](#) on Thu, 17 Jan 2008 19:26:23 GMT

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

Jonathan Corbet wrote:

>

> Heh, indeed. But we do seem to have a recurring problem of people

> wanting to extend sys_foo() beyond the confines of its original API.

> I've observed a few ways of doing that:

>

> - create sys_foo2() (or sys_foo64(), or sys_fooat(), or sys_pfoo(),

> or...) and add the new stuff there.

>

> The first approach has traditionally been the most popular. If we have

> a consensus that this is the way to extend system calls in the future,

> it would be nice to set that down somewhere. We could avoid a lot of

> API blind alleys that way.

>

I would argue it is the right approach. It lets the kernel system call entry dispatch directly to the system call for the "new" case, and to a compatibility thunk for the "old" case. It has the following desirable properties:

- No overhead for the "new" case.

- Minimal overhead for the "old" case.

- Easily dealt with by tools like strace that examine system calls.

-hpa

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>
