

Re: Syslets, Threadlets, generic AIO support, v6

Source: <http://linux.derkeiler.com/Mailing-Lists/Kernel/2007-05/msg13070.html>

- *From:* Ingo Molnar <mingo@xxxxxxx>
 - *Date:* Wed, 30 May 2007 10:42:52 +0200
-

* Ulrich Drepper <drepper@xxxxxxxxxxx> wrote:

Ingo Molnar wrote:

3 months ago i verified the published kevent vs. epoll benchmark and found that benchmark to be fatally flawed. When i redid it properly kevent showed no significant advantage over epoll.

I'm not going to judge your tests but saying there are no significant advantages is too one-sided. There is one huge advantage: the interface. A memory-based interface is simply the best form. File descriptors are a resource the runtime cannot transparently consume.

yeah – this is a fundamental design question for Linux i guess :-) glibc (and other infrastructure libraries) have a fundamental problem: they cannot (and do not) presently use persistent file descriptors to make use of kernel functionality, due to ABI side-effects. [applications can dup into an fd used by glibc, applications can close it – shells close fds blindly for example, etc.] Today glibc simply cannot open a file descriptor and keep it open while application code is running due to these problems.

we should perhaps enable glibc to have its separate fd namespace (or 'hidden' file descriptors at the upper end of the fd space) so that it can transparently listen to netlink events (or do epoll), without impacting the application fd namespace – instead of ducking to a memory based API as a workaround.

it is a serious flexibility issue that should not be ignored. The unified fd space is a blessing on one hand because it's simple and powerful, but it's also a curse because nested use of the fd space for libraries is currently not possible. But it should be detached from any fundamental question of kevent vs. epoll. (By improving library use of file descriptors we'll improve the utility of all syscalls – by ducking to a memory based API we only solve that particular event based usage.)

Re: Syslets, Threadlets, generic AIO support, v6

Ingo

–

To unsubscribe from this list: send the line "unsubscribe linux-kernel" in the body of a message to majordomo@xxxxxxxxxxxxxxxx

More majordomo info at <http://vger.kernel.org/majordomo-info.html>

Please read the FAQ at <http://www.tux.org/lkml/>