Subject: [PATCH 0/3] capabilities: per-process capbset
Posted by serue on Mon, 01 Oct 2007 14:40:31 GMT
View Forum Message <> Reply to Message

Here is a new per-process capability bounding set patchset
which I expect to send to linux-kernel soon. It makes
the capbset per-process. A process can only permanently
remove bits from it's bounding set, not add them. To
remove bits, CAP_SYS_ADMIN is currently needed. Maybe
that's not the best choice, but some privilege should
probably be required.

The intent is to allow a process tree to start with
certain capabilities, i.e. CAP_MKNOD, permanently
removed, so that running a setuid binary or one with
file capabilities will still not result in those
capabilities. The immediate use case for this is
containers/virtual servers.

I am not taking the task_capability_lock during
cap_prctl_setbset(), just as it is not taken when
capabilities are calculated during fork. That means
it can race with another task doing capsetp() on it,
and with capgetp(). I'm still looking for comments
on whether the fix I sent out last week is correct.
If it is, then I'll take the task_capability_lock
during cap_prctl_setbset().

thanks,
-serge

_____