

Re: mutex vs. local irqs (Was: 2.6.18 -mm merge plans)

Source: <http://linux.derkeiler.com/Mailing-Lists/Kernel/2006-06/msg02174.html>

- *From:* David Woodhouse <dwmw2@xxxxxxxxxxxxxx>
 - *Date:* Thu, 08 Jun 2006 12:01:06 +0100
-

On Thu, 2006-06-08 at 12:53 +0200, Ingo Molnar quoted Andrew:

Nonsense. `mutex_lock()` can sleep. Sleeping will enable interrupts. Therefore, hence, ergo ipso facto `mutex_lock()` can enable interrupts. QED, that's it.

But now, because some broken piece of hardware is coming out of reset/firmware asserting an interrupt we need to change the rules to be "`mutex_lock()` must preserve local interrupts if the lock is uncontended". Ditto `down()`, `down_read()` and `down_write()`.

And why does this bizarre restriction upon the implementation of our locking primitives exist? Because of your broken PIC and because of our inability to sort out the early boot code. And because the early boot code has this implicit knowledge that the locks will be uncontended, else we're toast.

We're doing `mutex_lock()`, `down()`, `down_read()` and `down_write()` with local interrupts disabled, which is a bug. We have explicit code in there to **disable** our runtime debugging checks because we know about this bug but don't know how to fix it.

I call that sucky.

OK, if you put it like that, and you're going to be consistent by declaring the disabling of `__might_sleep()` warnings to be sucky too, then I suppose we can buy that argument.

Yes, we need to sort out the early boot code. It isn't so much that we're unable as that nobody's really tried very hard. People seem scared of it -- they even invent pointless special cases like the 'earlyconsole' crap instead of just registering the damn consoles earlier, for example. `Register_console()` has `_always_` worked right from the beginning of `setup_arch()`, and I've often put it there.

Let's make a concerted effort to reorder the startup code so that we `_can_` enable interrupts and have slab working quite early. Ben's plans

Re: mutex vs. local irq (Was: 2.6.18 -mm merge plans)

for this look sane enough to me.

--

dwmw2

-

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/>