Re: [PATCH] don't use flush_tlb_all in suspend time

Source: http://linux.derkeiler.com/Mailing−Lists/Kernel/2006−04/msg06692.html

- *From*: Pavel Machek <pavel@xxxxxx>
- *Date*: Sun, 30 Apr 2006 14:04:22 +0200

On So 29−04−06 23:57:21, Andrew Morton wrote:

> Shaohua Li <shaohua.li@xxxxxxxxx> wrote:
>
> > On Sun, 2006−04−30 at 06:45 +0000, Pavel Machek wrote:
> >
> > > Hi!
> > >
> > > > flush_tlb_all uses on_each_cpu, which will
> > > > disable/enable interrupt.
> > > > In suspend/resume time, this will make
> > > > interrupt wrongly enabled.
> > > >
> > > > diff −puN
> > > > arch/i386/mm/init.c~flush_tlb_all_check
> > > > arch/i386/mm/init.c
> > > > −−−
> > > > linux−2.6.17−rc3/arch/i386/mm/init.c~flush_tlb_all_check
> > > > 2006−04−29 08:47:05.000000000 +0800
> > > > +++
> > > > linux−2.6.17−rc3−root/arch/i386/mm/init.c
> > > > 2006−04−29 08:48:15.000000000 +0800
> > > > @@ −420,7 +420,10 @@ void
> > > > zap_low_mappings (void)
> > > > #else
> > > > set_pgd(swapper_pg_dir+i, __pgd(0));
> > > > #endif
> > > > − flush_tlb_all();
> > > > + if (cpus_weight(cpu_online_map) == 1)
> > > > + local_flush_tlb();
> > > > + else
> > > > + flush_tlb_all();
> > > > }

> > Either it is okay to enable interrupts here −> unneccessary
> > and ugly
> > test, or it is not, and then we are broken in SMP case.
>
> It's not broken in SMP case, APs are offlined here in suspend/resume.

In which case, how's about this?

Certainly better, I'd say.

> @@ −420,7 +421,14 @@ void zap_low_mappings (void)
> #else
> set_pgd(swapper_pg_dir+i, __pgd(0));
> #endif
> − if (cpus_weight(cpu_online_map) == 1)
> + /*
> + * We can be called at suspend/resume time, with local interrupts
> + * disabled. But flush_tlb_all() requires that local interrupts be
> + * enabled.
> + *
> + * Happily, the APs are not yet started, so we can use local_flush_tlb() * in that case
> + */
> + if (num_online_cpus() == 1)
> local_flush_tlb();
> else
> flush_tlb_all();

But this still scares. It means calling convention is "may enable
interrupts with >1 cpu, may not with == 1 cpu".

Pavel
−−
Thanks for all the (sleeping) penguins.
−
To unsubscribe from this list: send the line "unsubscribe linux−kernel" in
the body of a message to majordomo@xxxxxxxxxxxxxxx
More majordomo info at http://vger.kernel.org/majordomo−info.html
Please read the FAQ at http://www.tux.org/lkml/