

# [patch] voluntary-preempt-2.6.9-rc1-bk4-Q7

**Source:** <http://linux.derkeiler.com/Mailing-Lists/Kernel/2004-09/0271.html>

---

**From:** Ingo Molnar ([mingo\\_at\\_elte.hu](mailto:mingo_at_elte.hu))

**Date:** 09/01/04

Date: Wed, 1 Sep 2004 15:51:22 +0200

To: [linux-kernel@vger.kernel.org](mailto:linux-kernel@vger.kernel.org)

i've released the -Q7 patch:

<http://redhat.com/~mingo/voluntary-preempt/voluntary-preempt-2.6.9-rc1-bk4-Q7>

ontop of:

<http://redhat.com/~mingo/voluntary-preempt/diff-bk-040828-2.6.8.1.bz2>

the main change in this patch are more SMP latency fixes. The stock kernel, even with CONFIG\_PREEMPT enabled, didnt have any spin-nicely preemption logic for the following, commonly used SMP locking primitives: read\_lock(), spin\_lock\_irqsave(), spin\_lock\_irq(), spin\_lock\_bh(), read\_lock\_irqsave(), read\_lock\_irq(), read\_lock\_bh(), write\_lock\_irqsave(), write\_lock\_irq(), write\_lock\_bh(). Only spin\_lock() and write\_lock() [the two simplest cases] where covered.

In addition to the preemption latency problems, the \_irq() variants in the above list didnt do any IRQ-enabling while spinning – possibly resulting in excessive irqsave sections of code!

-Q7 fixes all of these latency problems: we now re-enable interrupts while spinning in all possible cases, and a spinning op stays preemptible if this is a beginning of a new critical section.

there's also an SMP related tracing improvement in -Q7: the NMI tracing code now traces the other CPUs too – this way if an NMI hits a particular long section, we'll have a chance to see what the other CPU was doing. These show up as double do\_nmi() trace entries on a 2-CPU x86 box. The first one is the current CPU, subsequent entries are the other CPUs in the system.

(-Q7 is not that interesting to uniprocessor kernel users, but it would still be useful to test it, just to see nothing broke (on the compilation side), lots of spinlock code had to be changed.)

Ingo

—

To unsubscribe from this list: send the line "unsubscribe linux-kernel" in the body of a message to [majordomo@vger.kernel.org](mailto:majordomo@vger.kernel.org)

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

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