

Enabling RDPMC in user space by default

Source: <http://linux.derkeiler.com/Mailing-Lists/Kernel/2005-11/9419.html>

From: Andi Kleen (ak_at_suse.de)

Date: 11/29/05

Date: Tue, 29 Nov 2005 16:15:15 +0100
To: discuss@x86-64.org

Hallo,

I'm considering to enable CR4.PCE by default on x86-64/i386. Currently it's 0 which means RDPMC doesn't work. On x86-64 PMC 0 is always programmed to be a cycle counter, so it would be useful to be able to access this for measuring instructions. That's especially useful because RDTSC does not necessarily count cycles in the current P state (already the case on Intel CPUs and AMD's future direction seems to also to decouple it from cycles) Drawback is that it stops during idle, but that shouldn't be a big issue for normal measuring. It's not useful as a real timer anyways.

On Pentium 4 it also has the advantage that unlike RDTSC it's not serializing so should be much faster.

The kernel change would be to always set CR4.PCE to allow RDPMC in ring 3.

It would be actually a good idea to disable RDTSC in ring 3 too (because user space usually doesn't have enough information to make good use of it and gets it wrong), but I fear that will break too many applications right now.

Any comments on this?

–Andi

–

To unsubscribe from this list: send the line "unsubscribe linux-kernel" in the body of a message to 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/>