

Re: PATCH [PPC64]: dead processes never reaped

Source: <http://linux.derkeiler.com/Mailing-Lists/Kernel/2005-04/5104.html>

From: Benjamin Herrenschmidt (benh_at_kernel.crashing.org)

Date: 04/20/05

To: Linas Vepstas <linas@austin.ibm.com>

Date: Wed, 20 Apr 2005 15:44:10 +1000

On Mon, 2005-04-18 at 14:38 -0500, Linas Vepstas wrote:

>

> *Hi,*

>

> *The patch below appears to fix a problem where a number of dead processes
> linger on the system. On a highly loaded system, dozens of processes
> were found stuck in do_exit(), calling thier very last schedule(), and
> then being lost forever.*

Ok, we spent some time with Paul decrypting what `_switch_to` is supposed to do. Our understanding at this point is that the current code is correct on both ppc32 and ppc64, that is:

The "prev" passed in is always "current" and we don't see how it can be anything else. We use a local variable instead of current in the common code because accessing current can be slow on some architectures. I don't see any codepath where `prev != current` before `switch_to`.

If we didn't do some black magic that I explain below, `_switch_to` would switch the entire context, including stack, and thus including the value of "prev". Which means that we would always come back with prev beeing current, which is useless for reaping the old task. What we want is that this "prev" that was passed to `_switch_to()` is returned so that we can rip that previous task despite the change of context, that is basically prev has to be an invariant vs. the change of context in `switch_to`.

On ppc & ppc64, we implement that by passing that prev (or it's thread counterpart) to the assembly context switch code in r3. This code will preserve it and return it as-is (or re-transformed from thread to task).

So your problem must be somewhere else. I've looked at the `need_resched` code path and we always reload `prev = current` from a non-preemptible region, so it can't be wrong.

This was verified on 2.6.12-rc2, there might be something else wrong in an older kernel.

Linux-Kernel: Re: PATCH [PPC64]: dead processes never reaped

Ben.

—

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/>