On Sat, 3 May 2008 05:26:46 +0530
Balaji Rao <balajirrao@gmail.com> wrote:

> > yes, that would be good.
> OK, so when does account_system_time get called for the first time ? after
> IRQs are set up, is it ? So, where do we place the hook ?

Don't know - I'd need to dive in and work that out, and it's probably
better than you do this..

> Here's the patch.
>
> diff --git a/include/linux/percpu_counter.h b/include/linux/percpu_counter.h
> index 9007ccd..8a1b756 100644
> --- a/include/linux/percpu_counter.h
> +++ b/include/linux/percpu_counter.h
> @@ -21,7 +21,7 @@ struct percpu_counter {
> #ifdef CONFIG_HOTPLUG_CPU
>   struct list_head list; /* All percpu_counters are on a list */
> #endif
> - s32 *counters;
> + s32 counters[NR_CPUS];
> };

Please, no. That's a 4092-byte increase in sizeof(struct percpu_counter).
Hence a 12 kbyte increase in sizeof(struct ext3_sb_info). Let's just sort
out the cgroup startup ordering.

<looks at __percpu_alloc_mask>
<wanders off-topic>

Eric, is that optimal? alloc_percpu() will pass down cpu_possible_map in
`mask', and we only need to allocate enough slots to cover the
highest-set-bit in cpu_possible_map. However the implementation ignores
`mask' and does

    size_t sz = roundup(nr_cpu_ids * sizeof(void *), cache_line_size());
    void *pdata = kzalloc(sz, gfp);

Now, if the highest-set-bit in cpu_possible_map is always equal to
(1<<nr_cpu_ids) then it doesn't matter. But is that the case?

(If someone calls __percpu_alloc_mask with something that has less bits set
than cpu_possible_map then it surely is wasteful, but that sounds
unlikely).

_____
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers