

Re: [PATCH] Remove OOM killer from try_to_free_pages / all_unreclaimable braindamage

Source: <http://linux.derkeiler.com/Mailing-Lists/Kernel/2004-11/4927.html>

From: Chris Ross (chris_at_tebibyte.org)

Date: 11/18/04

Date: Thu, 18 Nov 2004 00:27:15 +0100
To: Werner Almesberger <wa@almesberger.net>

Werner Almesberger escreveu:

> *A process could declare itself as usual suspect. This would then be recorded as a per-task flag, to be inherited by children.*

I don't think this "I know I'm buggy, please kill me" flag is the right approach even if it can be made to work. The operating system has an overview of all the memory and can see when a particular process is basically making the machine unusable. It's quite likely that the process causing the trouble doesn't know (or hasn't admitted) that it's buggy and hasn't volunteered for early termination. As this means the kernel must be able to deal with a problematic process completely irrespective of whether it has set "kill me" flag or not the flag doesn't really buy you anything.

It is also specific to runaway processes that are clearly at fault. There is the related case where no particular process is faulty as such but the system as a whole can't cope with the demands being made.

On a related note, I would prefer to see victim processes who are not determined to be the cause of the trouble swapped out (i.e. **all** their pages pushed out to swap) and suspended (not allowed to run) as a first resort. The example I have in mind is on my machine when the daily cron run over commits causing standard daemons such as ntpd to be killed to make room. It would be preferable if the daemon was swapped out and just didn't run for minutes, or even hours if need be, but was allowed to run again once the system had settled down.

Of course, from recent discussion the system should not actually be killing off these daemons at all but that does seem to be resolved now. There are circumstances when there simply isn't enough RAM and swapping something out is preferable to killing it off. Of course, if there isn't sufficient swap space killing it should be the second resort. The last resort being panic.

So, the problem breaks down into three parts:

Linux-Kernel: Re: [PATCH] Remove OOM killer from try_to_free_pages / all_unreclaimable braindamage

- i) When should the oom killer be invoked.
- ii) How do we pick a victim process
- iii) How can we deal with the process in the most useful manner

Regards,
Chris R.

—

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/>