Subject: Re: [ckrm-tech] [RFC] [PATCH 0/3] Add group fairness to CFS
Posted by William Lee Irwin III on Thu, 31 May 2007 06:36:47 GMT
View Forum Message <> Reply to Message

On Wed, May 30, 2007 at 09:09:26PM -0700, William Lee Irwin III wrote:
>> It's not all that tricky.

On Thu, May 31, 2007 at 11:18:28AM +0530, Srivatsa Vaddagiri wrote:
> Hmm ..the fact that each task runs for a minimum of 1 tick seems to
> complicate the matters to me (when doing group fairness given a single
> level hierarchy). A user with 1000 (or more) tasks can be unduly
> advantaged compared to another user with just 1 (or fewer) task
> because of this?

Temporarily, yes. All this only works when averaged out. The basic idea is that you want a constant upper bound on the difference between the CPU time a task receives and the CPU time it was intended to get. This discretization is one of the larger sources of the "error" in the CPU time granted. The constant upper bound usually only applies to the largest difference for any task. When absolute values of differences are summed across tasks the aggregate will be O(tasks) because there's something almost like a constant per-task lower bound a la Heisenberg. It would have to get more exact the more tasks there are on the system for that to work, and something of the opposite actually holds.

It might be appropriate for the scheduler to dynamically adjust a periodic timer's period or to set up one-shot timers at involuntary preemption times in order to achieve more precise fairness in this sort of situation. In the case of few preemption points such one-shot code or low periodicity code would also save on taking interrupts that would otherwise manifest as overhead.

In short, a user with many tasks can reap a temporary advantage relative to users with fewer tasks because of this, but over time, longer-running tasks will receive the CPU time intended to within some constant upper bound, provided other things aren't broken.

-- wli

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