Subject: Re: [ckrm-tech] [RFC] [PATCH 0/3] Add group fairness to CFS
Posted by Peter Williams on Sat, 26 May 2007 00:17:42 GMT
View Forum Message <> Reply to Message

Srivatsa Vaddagiri wrote:
> Good example :) USER2's single task will have to share its CPU with
> USER1's 50 tasks (unless we modify the smpnice load balancer to
> disregard cpu affinity that is - which I would not prefer to do).

I don't think that ignoring cpu affinity is an option. Setting the cpu affinity of tasks is a deliberate policy action on the part of the system administrator and has to be honoured. Load balancing has to do the best it can in these circumstances which may mean sub optimal distribution of the load BUT it is result of a deliberate policy decision by the system administrator.

>
> Ingo/Peter, any thoughts here? CFS and smpnice probably is "broken"
> with respect to such example as above albeit for nice-based tasks.
>

See above. I think that faced with cpu affinity use by the system administrator that smpnice will tend towards a task to cpu allocation that is (close to) the best that can be achieved without violating the cpu affinity assignments. (It may take a little longer than normal but it should get there eventually.)

You have to assume that the system administrator knows what (s)he's doing and is willing to accept the impact of their policy decision on the overall system performance.

Having said that, if it was deemed necessary you could probably increase the speed at which the load balancer converged on a good result in the face of cpu affinity by keeping a "pinned weighted load" value for each run queue and using that to modify find_busiest_group() and find_busiest_queue() to be a bit smarter. But I'm not sure that it would be worth the added complexity.

Peter
--
Peter Williams pwil3058@bigpond.net.au

"Learning, n. The kind of ignorance distinguishing the studious."
-- Ambrose Bierce