Subject: Re: [PATCH 00/11] kmem controller for memcg: stripped down version
Posted by Glauber Costa on Wed, 27 Jun 2012 08:39:54 GMT
View Forum Message <> Reply to Message

On 06/27/2012 05:08 AM, David Rientjes wrote:
> On Tue, 26 Jun 2012, Andrew Morton wrote:
>
>> mm, maybe. Kernel developers tend to look at code from the point of
>> view "does it work as designed", "is it clean", "is it efficient", "do
>> I understand it", etc. We often forget to step back and really
>> consider whether or not it should be merged at all.
>>
>
> It's appropriate for true memory isolation so that applications cannot
> cause an excess of slab to be consumed. This allows other applications to
> have higher reservations without the risk of incurring a global oom
> condition as the result of the usage of other memcgs.

Just a note for Andrew, we we're in the same page: The slab cache limitation is not included in *this* particular series. The goal was always to have other kernel resources limited as well, and the general argument from David holds: we want a set of applications to run truly independently from others, without creating memory pressure on the global system.

The way history develop in this series, I started from the slab cache, and a page-level tracking appeared on that series. I then figured it would be better to start tracking something that is totally page-based, such as the stack - that already accounts for 70 % of the infrastructure, and then merge the slab code later. In this sense, it was just a strategy inversion. But both are, and were, in the goals.

> I'm not sure whether it would ever be appropriate to limit the amount of
> slab for an individual slab cache, however, instead of limiting the sum of
> all slab for a set of processes. With cache merging in slub this would
> seem to be difficult to do correctly.

Yes, I do agree.