Subject: Re: [PATCH v3 13/28] slub: create duplicate cache
Posted by Tejun Heo on Wed, 30 May 2012 08:02:35 GMT
View Forum Message <> Reply to Message

Hello, Glauber.

On Wed, May 30, 2012 at 4:54 PM, Glauber Costa <glommer@parallels.com> wrote:
> On 05/30/2012 05:29 AM, Tejun Heo wrote:
>>
>> The two goals for cgroup controllers that I think are important are
>> proper (no, not crazy perfect but good enough) isolation and an
>> implementation which doesn't impact !cg path in an intrusive manner -
>> if someone who doesn't care about cgroup but knows and wants to work
>> on the subsystem should be able to mostly ignore cgroup support. If
>> that means overhead for cgroup users, so be it.
>
>
> Well, my code in the slab is totally wrapped in static branches. They only
> come active when the first group is *limited* (not even created: you can
> have a thousand memcg, if none of them are kmem limited, nothing will
> happen).

Great, but I'm not sure why you're trying to emphasize that while my
point was about memory overhead and that it's OK to have some
overheads for cg users. :)

> After that, the cost paid is to find out at which cgroup the process is at.
> I believe that if we had a faster way for this (like for instance: if we had
> a single hierarchy, the scheduler could put this in a percpu variable after
> context switch - or any other method), then the cost of it could be really
> low, even when this is enabled.

Someday, hopefully.

> I will rework this series to try work more towards this goal, but at least
> for now I'll keep duplicating the caches. I still don't believe that a loose
> accounting to the extent Christoph proposed will achieve what we need this
> to achieve.

Yeah, I prefer your per-cg cache approach but do hope that it stays as
far from actual allocator code as possible. Christoph, would it be
acceptable if the cg logic is better separated?

Thanks.

--
tejun