Subject: Re: [PATCH v3 04/13] kmem accounting basic infrastructure
Posted by Glauber Costa on Wed, 26 Sep 2012 21:24:40 GMT
View Forum Message <> Reply to Message

On 09/27/2012 12:16 AM, Tejun Heo wrote:
> On Thu, Sep 27, 2012 at 12:02:14AM +0400, Glauber Costa wrote:
>> But think in terms of functionality: This thing here is a lot more
>> similar to swap than use_hierarchy. Would you argue that memsw should be
>> per-root ?
>
> I'm fairly sure you can make about the same argument about
> use_hierarchy. There is a choice to make here and one is simpler than
> the other. I want the additional complexity justified by actual use
> cases which isn't too much to ask for especially when the complexity
> is something visible to userland.
>
> So let's please stop arguing semantics. If this is definitely
> necessary for some use cases, sure let's have it. If not, let's
> consider it later. I'll stop responding on "inherent differences." I
> don't think we'll get anywhere with that.
>

If you stop responding, we are for sure not getting anywhere. I agree
with you here.

Let me point out one issue that you seem to be missing, and you respond
or not, your call.

"kmem_accounted" is not a switch. It is an internal representation only.
The semantics, that we discussed exhaustively in San Diego, is that a
group that is not limited is not accounted. This is simple and consistent.

Since the limits are still per-cgroup, you are actually proposing more
user-visible complexity than me, since you are adding yet another file,
with its own semantics.

About use cases, I've already responded: my containers use case is kmem
limited. There are people like Michal that specifically asked for
user-only semantics to be preserved. So your question for global vs
local switch (that again, doesn't exist; only a local *limit* exists)
should really be posed in the following way:
"Can two different use cases with different needs be hosted in the same
box?"

> Michal, Johannes, Kamezawa, what are your thoughts?
>

waiting! =)