
Subject: Re: [PATCH] fix bad behavior in use_hierarchy file
Posted by [Michal Hocko](#) on Tue, 26 Jun 2012 11:10:25 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Tue 26-06-12 14:31:51, Glauber Costa wrote:
> On 06/26/2012 11:56 AM, Michal Hocko wrote:
> >[Adding Ying to CC - they are using hierarchies AFAIU in their workloads]
> >
> >On Mon 25-06-12 13:49:08, Tejun Heo wrote:
> >[...]
> >>A bit of delta but is there any chance we can either deprecate
> >>.use_hierarchy or at least make it global toggle instead of subtree
> >>thing?
> >
> >So what you are proposing is to have all subtrees of the root either
> >hierarchical or not, right?
> >
> >>This seems needlessly complicated. :(
> >
> >Toggle wouldn't help much I am afraid. We would still have to
> >distinguish (non)hierarchical cases. And I am not sure we can make
> >everything hierarchical easily.
> >Most users (from my experience) ignored use_hierarchy for some reasons
> >and the end results might be really unexpected for them if they used
> >deeper subtrees (which might be needed due to combination with other
> >controller(s)).
> >
> Do we have any idea about who those users are, and how is their
> setup commonly done?

Well, most of them use memory controller with combination of other
controller - usually cpuset or cpu - and memcg is used to cap the amount
of memory for each respective group. As I said most of those users
were not aware of use_hierarchy at all.

> We can propose work arounds here, but not without first knowing work
> arounds to what =p

No, please no workarounds. It will be even bigger mess.
Maybe a global switch is the first step in the right direction (on by
default). If somebody encounters any issue we can say it can be turned
off (something like one time switch) or advise on how to fix their
layout to fit hierarchy better. We can put WARN_ON_ONCE when the knob is
set to 0 in the second stage and finally remove the whole knob.

> One thing that would really influence this, for instance, is whether
> or not they limit at all levels in the tree, etc.

Yes and independently. But it's true that I haven't seen many of them to be honest, people usually use a flat structures.

--

Michal Hocko
SUSE Labs
SUSE LINUX s.r.o.
Lihovarska 1060/12
190 00 Praha 9
Czech Republic
