

---

Subject: Re: [RFC 2/4] memcg: make it suck faster  
Posted by [Daniel P. Berrange](#) on Wed, 26 Sep 2012 09:03:43 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

On Wed, Sep 26, 2012 at 12:53:21PM +0400, Glauber Costa wrote:

> On 09/26/2012 01:02 AM, Andrew Morton wrote:

> >> nomemcg : memcg compile disabled.

> >> > base : memcg enabled, patch not applied.

> >> > bypassed : memcg enabled, with patch applied.

> >> >

> >> > base bypassed

> >> > User 109.12 105.64

> >> > System 1646.84 1597.98

> >> > Elapsed 229.56 215.76

> >> >

> >> > nomemcg bypassed

> >> > User 104.35 105.64

> >> > System 1578.19 1597.98

> >> > Elapsed 212.33 215.76

> >> >

> >> > So as one can see, the difference between base and nomemcg in terms

> >> > of both system time and elapsed time is quite drastic, and consistent

> >> > with the figures shown by Mel Gorman in the Kernel summit. This is a

> >> > ~ 7 % drop in performance, just by having memcg enabled. memcg functions

> >> > appear heavily in the profiles, even if all tasks lives in the root

> >> > memcg.

> >> >

> >> > With bypassed kernel, we drop this down to 1.5 %, which starts to fall

> >> > in the acceptable range. More investigation is needed to see if we can

> >> > claim that last percent back, but I believe at last part of it should

> >> > be.

> > Well that's encouraging. I wonder how many users will actually benefit

> > from this - did I hear that major distros are now using memcg in some

> > system-infrastructure-style code?

> >

>

> If they do, they actually be come "users of memcg". This here is aimed

> at non-users of memcg, which given all the whining about it, it seems to

> be plenty.

>

> Also, I noticed, for instance, that libvirt is now creating memcg

> hierarchies for lxc and qemu as placeholders, before you actually create

> any vm or container.

This is mostly just lazyness on our part. There's no technical reason why we can't delay creating our intermediate cgroups until we actually have a VM ready to start, it was just simpler to create them when we started the main daemon.

Daniel

--

|: <http://berrange.com>    -o-    <http://www.flickr.com/photos/dberrange/> :|  
|: <http://libvirt.org>        -o-        <http://virt-manager.org> :|  
|: <http://autobuild.org>    -o-        <http://search.cpan.org/~danberr/> :|  
|: <http://entangle-photo.org>    -o-        <http://live.gnome.org/gtk-vnc> :|

---