Subject: Re: [PATCH v3 04/13] kmem accounting basic infrastructure
Posted by Tejun Heo on Sun, 30 Sep 2012 08:02:49 GMT

View Forum Message <> Reply to Message

On Sun, Sep 30, 2012 at 04:57:00PM +0900, Tejun Heo wrote:
> On Thu, Sep 27, 2012 at 10:45:01PM +0400, Glauber Costa wrote:
> > > Can you please give other examples of cases where this type of issue
> > > exists (plenty of shared kernel data structure which is inherent to
> > > the workload at hand)? Until now, this has been the only example for
> > > this type of issues.
> >
> > Yes. the namespace related caches (*), all kinds of sockets and network
> > structures, other file system structures like file struct, vm areas, and
> > pretty much everything a full container does.
> >
> > (*) we run full userspace, so we have namespaces + cgroups combination.
>
> This is probably me being dumb but wouldn't resources used by full
> namespaces be mostly independent? Which parts get shared? Also, if
> you do full namespace, isn't it more likely that you would want fuller
> resource isolation too?

Just a thought about dentry/inode. Would it make sense to count total
number of references per cgroup and charge the total amount according
to that? Reference counts are how the shared ownership is represented
after all. Counting total per cgroup isn't accurate and pathological
cases could be weird tho.

Thanks.

--
tejun