
Subject: Re: [ckrm-tech] [PATCH] BC: resource
beancounters (v4) (added user memory)

Posted by [Chandra Seetharaman](#) on Wed, 13 Sep 2006 01:10:18 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Tue, 2006-09-12 at 17:39 -0700, Rohit Seth wrote:

<snip>

> > yes, it would be there, but is not heavy, IMO.

>

> I think anything greater than 1% could be a concern for people who are

> not very interested in containers but would be forced to live with them.

If they are not interested in resource management and/or containers, i
do not think they need to pay.

>

> > >

> > > >

> > > > And anything running outside a container should be limited by default

> > > > Linux settings.

> > > >

> > > > note that the resource available to the default RG will be (total system

> > > > resource - allocated to RGs).

> > >

> > > I think it will be preferable to not change the existing behavior for

> > > applications that are running outside any container (in your case

> > > default resource group).

> >

> > hmm, when you provide QoS for a set of apps, you will affect (the

> > resource availability of) other apps. I don't see any way around it. Any

> > ideas ?

>

> When I say, existing behavior, I mean not getting impacted by some

> artificial limits that are imposed by container subsystem. IOW, if a

That is what I understood and replied above.

> sysadmin is okay to have certain apps running outside of container then

> he is basically forgoing any QoS for any container on that system.

Not at all. If the container they are interested in is guaranteed, I do
not see how apps running outside a container would affect them.

<snip>

> > > > Not really.

> > > > - Each RG will have a guarantee and limit of each resource.

> > > > - default RG will have (system resource - sum of guarantees)

> > > > - Every RG will be guaranteed some amount of resource to provide QoS

> > > > - Every RG will be limited at "limit" to prevent DoS attacks.

> > > > - Whoever doesn't care either of those set them to don't care values.

> > > >
> > >
> > > For the cases that put this don't care, do you depend on existing
> > > reclaim algorithm (for memory) in kernel?
> >
> > Yes.
>
> So one container with these don't care condition(s) can turn the whole
> guarantee thing bad. Because existing kernel reclaimer does not know
> about memory commitments to other containers. Right?

No, the reclaimer would free up pages associated with the don't care RGs
(as the user don't care about the resource made available to them).

<snip>

> > > If the limits are set appropriately so that containers total memory
> > > consumption does not exceed the system memory then there shouldn't be
> > > any QoS issue (to whatever extent it is applicable for specific
> > > scenario).
> >
> > Then you will not be work-conserving (IOW over-committing), which is one
> > of the main advantage of this type of feature.
> >
>
> If for the systems where QoS is important, not over-committing will be
> fine (at least to start with).

The problem is that you can't do it with just limit.

--

Chandra Seetharaman | Be careful what you choose....
- sekharan@us.ibm.com |you may get it.
