Subject: Re: [ckrm-tech] [PATCH] BC: resource beancounters (v4) (added user memory)
Posted by Rohit Seth on Wed, 13 Sep 2006 01:33:38 GMT
View Forum Message <> Reply to Message

On Tue, 2006-09-12 at 18:13 -0700, Chandra Seetharaman wrote:
> On Tue, 2006-09-12 at 17:43 -0700, Rohit Seth wrote:
> <snip>
>
> > > It won't be a complete solution, as the user won't be able to
> > > - set both guarantee and limit for a resource group
> > > - use limit on some and guarantee on some
> > > - optimize the usage of available resources
> >
> > I think, if we have some of the dynamic resource limit adjustments
> > possible then some of the above functionality could be achieved. And I
> > think that could be a good start point.
>
>
> Yes, dynamic resource adjustments should be available. But, you can't
> expect the sysadmin to sit around and keep tweaking the limits so as to
> achieve the QoS he wants. (Even if you have an application sitting and
> doing it, as I pointed in other email it may not be possible for
> different scenarios).
> >

As said earlier, if strict QoS is desired then system should be appropriately partitioned so that the sum of limits doesn't exceed physical limit (that is cost of QoS). Let us first get at least that much settled on and accepted in mainline before getting into these esoteric features.

-rohit

Subject: Re: [ckrm-tech] [PATCH] BC: resource beancounters (v4) (added user memory)
Posted by Chandra Seetharaman on Wed, 13 Sep 2006 22:24:07 GMT
View Forum Message <> Reply to Message

On Tue, 2006-09-12 at 18:33 -0700, Rohit Seth wrote:
> On Tue, 2006-09-12 at 18:13 -0700, Chandra Seetharaman wrote:
> > On Tue, 2006-09-12 at 17:43 -0700, Rohit Seth wrote:
> > <snip>
> >
> > > > It won't be a complete solution, as the user won't be able to
> > > > - set both guarantee and limit for a resource group

> > > > - use limit on some and guarantee on some
> > > > - optimize the usage of available resources
> > >
> > > I think, if we have some of the dynamic resource limit adjustments
> > > possible then some of the above functionality could be achieved. And I
> > > think that could be a good start point.
> >
> >
> > Yes, dynamic resource adjustments should be available. But, you can't
> > expect the sysadmin to sit around and keep tweaking the limits so as to
> > achieve the QoS he wants. (Even if you have an application sitting and
> > doing it, as I pointed in other email it may not be possible for
> > different scenarios).
> > >
>
>
> As said earlier, if strict QoS is desired then system should be
> appropriately partitioned so that the sum of limits doesn't exceed
> physical limit (that is cost of QoS). Let us first get at least that
> much settled on and accepted in mainline before getting into these
> esoteric features.
>

esoteric ?! Please look at the different operating system that provide
resource management and other resource management capability providers.
All of them have both guarantees and limits (they might call them
differently).

Here are a few:
http://www.hp.com/go/prm
http://www.sun.com/software/resourcemgr/
http://www.redbooks.ibm.com/redbooks/pdfs/sg245977.pdf
http://www.vmware.com/pdf/vmware_drs_wp.pdf
http://www.aurema.com

--

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

Subject: Re: [ckrm-tech] [PATCH] BC: resource
beancounters (v4) (added user memory)
Posted by Rohit Seth on Thu, 14 Sep 2006 01:27:12 GMT

On Wed, 2006-09-13 at 15:24 -0700, Chandra Seetharaman wrote:
> On Tue, 2006-09-12 at 18:33 -0700, Rohit Seth wrote:
> > On Tue, 2006-09-12 at 18:13 -0700, Chandra Seetharaman wrote:
> > > On Tue, 2006-09-12 at 17:43 -0700, Rohit Seth wrote:
> > > <snip>
> > >
> > > > > It won't be a complete solution, as the user won't be able to
> > > > > - set both guarantee and limit for a resource group
> > > > > - use limit on some and guarantee on some
> > > > > - optimize the usage of available resources
> > > >
> > > > I think, if we have some of the dynamic resource limit adjustments
> > > > possible then some of the above functionality could be achieved. And I
> > > > think that could be a good start point.
> > >
> > >
> > > Yes, dynamic resource adjustments should be available. But, you can't
> > > expect the sysadmin to sit around and keep tweaking the limits so as to
> > > achieve the QoS he wants. (Even if you have an application sitting and
> > > doing it, as I pointed in other email it may not be possible for
> > > different scenarios).
> > >
> >
> > As said earlier, if strict QoS is desired then system should be
> > appropriately partitioned so that the sum of limits doesn't exceed
> > physical limit (that is cost of QoS). Let us first get at least that
> > much settled on and accepted in mainline before getting into these
> > esoteric features.
> >
>
> esoteric ?! Please look at the different operating system that provide
> resource management and other resource management capability providers.
> All of them have both guarantees and limits (they might call them
> differently).
>

Is this among the very first features you would like (to get absolutely
right) before containers get in mm tree? Or is this something that can
wait after the minimal infrastructure is in Andrew's tree and the code
gets wider testing...And above all we have agreed upon user interface.

-rohit

Subject: Re: [ckrm-tech] [PATCH] BC:

resource beancounters (v4) (added user memory)
Posted by Chandra Seetharaman on Thu, 14 Sep 2006 23:28:15 GMT
View Forum Message <> Reply to Message

On Wed, 2006-09-13 at 18:27 -0700, Rohit Seth wrote:

<snip>
> > > As said earlier, if strict QoS is desired then system should be
> > > appropriately partitioned so that the sum of limits doesn't exceed
> > > physical limit (that is cost of QoS). Let us first get at least that
> > > much settled on and accepted in mainline before getting into these
> > > esoteric features.
> > >
> >
> > esoteric ?! Please look at the different operating system that provide
> > resource management and other resource management capability providers.
> > All of them have both guarantees and limits (they might call them
> > differently).
> >
>
> Is this among the very first features you would like (to get absolutely
> right) before containers get in mm tree? Or is this something that can

Let me make it clear, I am interested in resource management and not in containers.

IMO, for resource management to work as expected (as is in other OSes), guarantee is needed. It will be a good idea to have it from start as it would affect the design of controllers.

For example, instead of writing two controllers (one to control limit and another to provide guarantee), controller writers can provide both in a single controller. (OpenVZ has two parameters, oomguarpages and vmguarpages whose purpose is to provide some sort of guarantee using the barrier and/or limit available in BC)

> wait after the minimal infrastructure is in Andrew's tree and the code
> gets wider testing...And above all we have agreed upon user interface.
>
> -rohit
>

--

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

Subject: Re: [ckrm-tech] [PATCH] BC: resource beancounters (v4) (added user memory)
Posted by dev on Fri, 15 Sep 2006 09:22:09 GMT
View Forum Message <> Reply to Message

Chandra Seetharaman wrote:

> esoteric ?! Please look at the different operating system that provide
> resource management and other resource management capability providers.
> All of them have both guarantees and limits (they might call them
> differently).
>
> Here are a few:
> http://www.hp.com/go/prm
> http://www.sun.com/software/resourcemgr/
> http://www.redbooks.ibm.com/redbooks/pdfs/sg245977.pdf
> http://www.vmware.com/pdf/vmware_drs_wp.pdf
> http://www.aurema.com
have you ever tested any of these?!
there is no _memory_ guarantees AFAIK in all of them except
for VMware which can reserve required amount of RAM for VM.

Kirill

Subject: Re: [ckrm-tech] [PATCH] BC: resource beancounters (v4) (added user memory)
Posted by Rohit Seth on Fri, 15 Sep 2006 16:52:40 GMT
View Forum Message <> Reply to Message

On Fri, 2006-09-15 at 13:26 +0400, Kirill Korotaev wrote:
> Chandra Seetharaman wrote:
>
> > esoteric ?! Please look at the different operating system that provide
> > resource management and other resource management capability providers.
> > All of them have both guarantees and limits (they might call them
> > differently).
> >
> > Here are a few:
> > http://www.hp.com/go/prm
> > http://www.sun.com/software/resourcemgr/
> > http://www.redbooks.ibm.com/redbooks/pdfs/sg245977.pdf
> > http://www.vmware.com/pdf/vmware_drs_wp.pdf
> > http://www.aurema.com
> have you ever tested any of these?!
> there is no _memory_ guarantees AFAIK in all of them except

Interesting...

> for VMware which can reserve required amount of RAM for VM.

It is much easier to provide guarantees in complete virtual
environments. But then you pay the cost in terms of performance.

I think we should punt on hard guarantees and fractions for the first
draft. Keep the implementation simple.

-rohit

Subject: Re: Re: [ckrm-tech] [PATCH] BC: resource beancounters (v4) (added user memory)
Posted by kir on Fri, 15 Sep 2006 21:20:27 GMT
View Forum Message <> Reply to Message

Rohit Seth wrote:
> On Fri, 2006-09-15 at 13:26 +0400, Kirill Korotaev wrote:
>
> <...skipped...>
>
>> for VMware which can reserve required amount of RAM for VM.
>>
>
> It is much easier to provide guarantees in complete virtual
> environments. But then you pay the cost in terms of performance.
>
"Complete virtual environments" vs. "contaners" is not [only] about
performance! In the end, given a proper set of dirty and no-so-dirty
hacks in software and hardware, their performance will be close to native.

Containers vs. other virtualization types is more about utilization,
density, scalability, portability.

Speaking of guarantees, yes, guarantees is easy, you just reserve such
amount of RAM for your VM and that is all. But the fact is usually some
part of that RAM will not be utilized by this particular VM. But since
it is reserved, it can not be utilized by other VMs -- and we end up
just wasting some resources. Containers, given a proper resource
management and configuration, can have some guarantees and still be able
to utilize all the RAM available in the system. This difference can be
metaphorically expressed as a house divided into rooms. Dividing walls
can either be hard or flexible. With flexible walls, room (container)
owner have a guarantee of minimal space in your room, but if a few
guests come for a moment, the walls can move to make more space (up to
the limit). So the flexibility is measured as the delta between a
guarantee and a limit.

This flexibility leads to higher utilization, and this flexibility is
one of the reasons for better density (a few times higher than that of a
paravirtualization solution).

I will not touch scalability and portability topics here to make things
simpler.
> I think we should punt on hard guarantees and fractions for the first
> draft. Keep the implementation simple.
>
Do I understand it right that with hard guarantees we loose the
flexibility I have just described? If this is the case, I do not like it.

Subject: Re: Re: [ckrm-tech] [PATCH] BC: resource beancounters (v4) (added user memory)
Posted by Rohit Seth on Fri, 15 Sep 2006 21:58:53 GMT
View Forum Message <> Reply to Message

On Sat, 2006-09-16 at 01:21 +0400, Kir Kolyshkin wrote:
> Rohit Seth wrote:
> > On Fri, 2006-09-15 at 13:26 +0400, Kirill Korotaev wrote:
> >
> > <...skipped...>
> >
> >> for VMware which can reserve required amount of RAM for VM.
> >>
> >
> > It is much easier to provide guarantees in complete virtual
> > environments. But then you pay the cost in terms of performance.
> >
> "Complete virtual environments" vs. "contaners" is not [only] about
> performance! In the end, given a proper set of dirty and no-so-dirty
> hacks in software and hardware, their performance will be close to native.
>

I don't think there is current generation of Virtualization HW/SW that
can live with o-2% performance loss for all workloads (like the way
containers do).

> Containers vs. other virtualization types is more about utilization,
> density, scalability, portability.
>

I agree with most of it (except portability as using latest HW
technologies you can run unmodified guests in virtualized environment).

> Speaking of guarantees, yes, guarantees is easy, you just reserve such

> amount of RAM for your VM and that is all. But the fact is usually some
> part of that RAM will not be utilized by this particular VM. But since
> it is reserved, it can not be utilized by other VMs -- and we end up
> just wasting some resources. Containers, given a proper resource
> management and configuration, can have some guarantees and still be able
> to utilize all the RAM available in the system. This difference can be
> metaphorically expressed as a house divided into rooms. Dividing walls
> can either be hard or flexible. With flexible walls, room (container)
> owner have a guarantee of minimal space in your room, but if a few
> guests come for a moment, the walls can move to make more space (up to
> the limit). So the flexibility is measured as the delta between a
> guarantee and a limit.
>
> This flexibility leads to higher utilization, and this flexibility is
> one of the reasons for better density (a few times higher than that of a
> paravirtualization solution).
>

I guess as far as memory is concerned, virtualized solutions can also
techniques like ballooning to oversubscribe memory. But I agree that we
will almost always be able to pack things tighter in container
environment.

> I will not touch scalability and portability topics here to make things
> simpler.
> > I think we should punt on hard guarantees and fractions for the first
> > draft. Keep the implementation simple.
> >
> Do I understand it right that with hard guarantees we loose the
> flexibility I have just described? If this is the case, I do not like it.

With hard guarantees, you will also end up making hooks in generic part
of kernel which could be considered invasive. And yes, if you are
making a hard guarantee then you will some how make sure that amount of
resource is available all the time for that container. As you mentioned
this is not the most optimal use of resources. And that is why I don't
want to incorporate that in at least the first draft. Please look at
the container kernel patches that I sent out yesterday. They allow the
containers to go over board with memory as long as there is no pressure.
But the moment there is any pressure on memory, pages belonging to over
the limit containers get freed or swapped first.

-rohit

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

Posted by Chandra Seetharaman on Mon, 18 Sep 2006 23:59:11 GMT
View Forum Message <> Reply to Message

On Fri, 2006-09-15 at 13:26 +0400, Kirill Korotaev wrote:
> Chandra Seetharaman wrote:
>
> > esoteric ?! Please look at the different operating system that provide
> > resource management and other resource management capability providers.
> > All of them have both guarantees and limits (they might call them
> > differently).
> >
> > Here are a few:
> > http://www.hp.com/go/prm
> > http://www.sun.com/software/resourcemgr/
> > http://www.redbooks.ibm.com/redbooks/pdfs/sg245977.pdf
> > http://www.vmware.com/pdf/vmware_drs_wp.pdf
> > http://www.aurema.com
> have you ever tested any of these?!
> there is no _memory_ guarantees AFAIK in all of them except
> for VMware which can reserve required amount of RAM for VM.

I have tried VMware but no others. Nevertheless, I was not talking only
in the context of memory, I was talking about the features
infrastructure should provide (for different resource controllers).
>
> Kirill
>
>
> ---------- ----------
> Using Tomcat but need to do more? Need to support web services, security?
> Get stuff done quickly with pre-integrated technology to make your job easier
> Download IBM WebSphere Application Server v.1.0.1 based on Apache Geronimo
> http://sel.as-us.falkag.net/sel?cmd=lnk&kid=120709&b id=263057&dat=121642
> _____
> ckrm-tech mailing list
> https://lists.sourceforge.net/lists/listinfo/ckrm-tech
--

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

Subject: Re: [ckrm-tech] Re: [PATCH] BC: resource beancounters (v4) (added user memory)
Posted by Chandra Seetharaman on Tue, 19 Sep 2006 00:02:27 GMT
View Forum Message <> Reply to Message

On Sat, 2006-09-16 at 01:21 +0400, Kir Kolyshkin wrote:
> Rohit Seth wrote:
> > On Fri, 2006-09-15 at 13:26 +0400, Kirill Korotaev wrote:
> >
> > <...skipped...>
> >
> >> for VMware which can reserve required amount of RAM for VM.
> >>
> >
> > It is much easier to provide guarantees in complete virtual
> > environments. But then you pay the cost in terms of performance.
> >
> "Complete virtual environments" vs. "contaners" is not [only] about
> performance! In the end, given a proper set of dirty and no-so-dirty
> hacks in software and hardware, their performance will be close to native.
>
> Containers vs. other virtualization types is more about utilization,
> density, scalability, portability.
>
> Speaking of guarantees, yes, guarantees is easy, you just reserve such
> amount of RAM for your VM and that is all. But the fact is usually some
> part of that RAM will not be utilized by this particular VM. But since
> it is reserved, it can not be utilized by other VMs -- and we end up
> just wasting some resources. Containers, given a proper resource
> management and configuration, can have some guarantees and still be able
> to utilize all the RAM available in the system. This difference can be
> metaphorically expressed as a house divided into rooms. Dividing walls
> can either be hard or flexible. With flexible walls, room (container)
> owner have a guarantee of minimal space in your room, but if a few
> guests come for a moment, the walls can move to make more space (up to
> the limit). So the flexibility is measured as the delta between a
> guarantee and a limit.
>
> This flexibility leads to higher utilization, and this flexibility is
> one of the reasons for better density (a few times higher than that of a
> paravirtualization solution).
>
> I will not touch scalability and portability topics here to make things
> simpler.
> > I think we should punt on hard guarantees and fractions for the first
> > draft. Keep the implementation simple.
> >
> Do I understand it right that with hard guarantees we loose the
> flexibility I have just described? If this is the case, I do not like it.
>

If I understand your description correctly (describing flexibility to be
the ability to move the resource usage between guarantee and limit),

then NO, you will not loose that flexibility.

> ---------- ----------
> Using Tomcat but need to do more? Need to support web services, security?
> Get stuff done quickly with pre-integrated technology to make your job easier
> Download IBM WebSphere Application Server v.1.0.1 based on Apache Geronimo
> http://sel.as-us.falkag.net/sel?cmd=lnk&kid=120709&b id=263057&dat=121642
> _____
> ckrm-tech mailing list
> https://lists.sourceforge.net/lists/listinfo/ckrm-tech
--

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