Subject: Re: [dm-devel] [PATCH 0/2] dm-band: The I/O bandwidth controller: Overview
Posted by Alasdair G Kergon on Wed, 23 Jan 2008 14:47:17 GMT
View Forum Message <> Reply to Message

On Wed, Jan 23, 2008 at 09:53:50PM +0900, Ryo Tsuruta wrote:
> Dm-band gives bandwidth to each job according to its weight,
> which each job can set its own value to.
> At this time, a job is a group of processes with the same pid or pgrp or uid.

It seems to rely on 'current' to classify bios and doesn't do it until the map
function is called, possibly in a different process context, so it won't
always identify the original source of the I/O correctly: people need to take
this into account when designing their group configuration and so this should
be mentioned in the documentation.

I've uploaded it here while we consider ways we might refine the architecture and
interfaces etc.:

  http://www.kernel.org/pub/linux/kernel/people/agk/patches/2.6/editing/dm-add-band-target.patch

Alasdair
--
agk@redhat.com
_____
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers

Subject: Re: [dm-devel] [PATCH 0/2] dm-band: The I/O bandwidth controller: Overview
Posted by Hirokazu Takahashi on Wed, 23 Jan 2008 16:21:03 GMT
View Forum Message <> Reply to Message

Hi,

> On Wed, Jan 23, 2008 at 09:53:50PM +0900, Ryo Tsuruta wrote:
> > Dm-band gives bandwidth to each job according to its weight,
> > which each job can set its own value to.
> > At this time, a job is a group of processes with the same pid or pgrp or uid.
>
> It seems to rely on 'current' to classify bios and doesn't do it until the map
> function is called, possibly in a different process context, so it won't
> always identify the original source of the I/O correctly:

Yes, this should be mentioned in the document with the current implementation
as you pointed out.

By the way, I think once a memory controller of cgroup is introduced, it will
help to track down which cgroup is the original source.

> people need to take
> this into account when designing their group configuration and so this should
> be mentioned in the documentation.
>
> I've uploaded it here while we consider ways we might refine the architecture and
> interfaces etc.:
>
>
http://www.kernel.org/pub/linux/kernel/people/agk/patches/2.6/editing/dm-add-band-target.patch
>
> Alasdair
> --
> agk@redhat.com

Thank you,
Hirokazu Takahashi.

_____

Subject: Re: [dm-devel] [PATCH 0/2] dm-band: The I/O bandwidth controller:
Overview
Posted by yamamoto on Thu, 24 Jan 2008 03:38:07 GMT
View Forum Message <> Reply to Message

> Hi,
>
> > On Wed, Jan 23, 2008 at 09:53:50PM +0900, Ryo Tsuruta wrote:
> > > Dm-band gives bandwidth to each job according to its weight,
> > > which each job can set its own value to.
> > > At this time, a job is a group of processes with the same pid or pgrp or uid.
> >
> > It seems to rely on 'current' to classify bios and doesn't do it until the map
> > function is called, possibly in a different process context, so it won't
> > always identify the original source of the I/O correctly:
>
> Yes, this should be mentioned in the document with the current implementation
> as you pointed out.
>
> By the way, I think once a memory controller of cgroup is introduced, it will
> help to track down which cgroup is the original source.

do you mean to make this a part of the memory subsystem?

YAMAMOTO Takashi

_____

Subject: Re: [dm-devel] [PATCH 0/2] dm-band: The I/O bandwidth controller: Overview
Posted by Hirokazu Takahashi on Thu, 24 Jan 2008 10:14:56 GMT
View Forum Message <> Reply to Message

Hi,

> > > On Wed, Jan 23, 2008 at 09:53:50PM +0900, Ryo Tsuruta wrote:
> > > > Dm-band gives bandwidth to each job according to its weight,
> > > > which each job can set its own value to.
> > > > At this time, a job is a group of processes with the same pid or pgrp or uid.
> > >
> > > It seems to rely on 'current' to classify bios and doesn't do it until the map
> > > function is called, possibly in a different process context, so it won't
> > > always identify the original source of the I/O correctly:
> >
> > Yes, this should be mentioned in the document with the current implementation
> > as you pointed out.
> >
> > By the way, I think once a memory controller of cgroup is introduced, it will
> > help to track down which cgroup is the original source.
>
> do you mean to make this a part of the memory subsystem?

I just think if the memory subsystem is in front of us, we don't need to
reinvent the wheel.

But I don't have a concrete image how the interface between dm-band and
the memory subsystem should be designed yet. I'd be appreciate if some of
the cgroup developers give some ideas about it.

Thanks,
Hirokazu Takahashi.

> YAMAMOTO Takashi

_____

Subject: Re: [dm-devel] [PATCH 0/2] dm-band: The I/O bandwidth controller: Overview
Posted by yamamoto on Fri, 25 Jan 2008 06:26:41 GMT
View Forum Message <> Reply to Message

> Hi,
>
> > > > On Wed, Jan 23, 2008 at 09:53:50PM +0900, Ryo Tsuruta wrote:
> > > > > Dm-band gives bandwidth to each job according to its weight,
> > > > > which each job can set its own value to.
> > > > > At this time, a job is a group of processes with the same pid or pgrp or uid.
> > > >
> > > > It seems to rely on 'current' to classify bios and doesn't do it until the map
> > > > function is called, possibly in a different process context, so it won't
> > > > always identify the original source of the I/O correctly:
> > >
> > > Yes, this should be mentioned in the document with the current implementation
> > > as you pointed out.
> > >
> > > By the way, I think once a memory controller of cgroup is introduced, it will
> > > help to track down which cgroup is the original source.
> >
> > do you mean to make this a part of the memory subsystem?
>
> I just think if the memory subsystem is in front of us, we don't need to
> reinvent the wheel.
>
> But I don't have a concrete image how the interface between dm-band and
> the memory subsystem should be designed yet. I'd be appreciate if some of
> the cgroup developers give some ideas about it.

the current implementation of memory subsystem associates pages to
cgroups directly, rather than via tasks. so it isn't straightforward to
use the information for other classification mechanisms like yours which
might not share the view of "hierarchy" with the memory subsystem.

YAMAMOTO Takashi

>
> Thanks,
> Hirokazu Takahashi.
>
>
> > YAMAMOTO Takashi