Subject: [PATCH O/4] Block I/O tracking
Posted by Hirokazu Takahashi on Tue, 18 Mar 2008 09:22:51 GMT
View Forum Message <> Reply to Message

Hi,

When you want to implement some kind of Block I/O controllers, you have
to determine who issued each I/O. I just implemented this feature,
with which you can track down the I/Os.

When you have to find the owner which issued the I/O, it is the one which
owns the page where the IO is going to start. The cgroup memory subsystem
already has this feature, so I realized that it would make easy to
implemented Block I/O tracking mechanism on the memory subsystem.
I named it "bio cgroup."

I made dm-ioband -- I/O bandwidth controller -- work with the bio cgroup,
whose implementation is just experimental though.

I have a plan on making the bio cgroup support io_context. Each bio
cgroup will have one or more io_contexts so the I/O bandwidth controller
can use it to control the bandwidths.
I also have another plan on move the implementation of dm-ioband from
the device mapper layer to somewhere before the I/O schedulers
in the block layer.

The following patches are against linux-2.6.25-rc5-mm1 and you have to
apply the patch of dm-ioband v0.0.3, which you can download from
http://people.valinux.co.jp/~ryov/dm-ioband/patches/dm-ioband-0.0.3.patch
before applying the following patches.

Let's say you want make two bio cgroups and assign them to ioband
device "ioband1". First, you have to mount the bio cgroup filesystem.

```
 # mount -t cgroup -o bio none /cgroup/bio
```

Then, you make new bio cgroups and put some processes in them.

```
 # mkdir /cgroup/bio/bgroup1
 # mkdir /cgroup/bio/bgroup2
 # echo 1234 /cgroup/bio/bgroup1/tasks
 # echo 5678 /cgroup/bio/bgroup1/tasks
```

Now you check the ids of the bio cgroups which you just created.

```
 # cat /cgroup/bio/bgroup1/bio.id
   1
 # cat /cgroup/bio/bgroup2/bio.id
```

2

Finally, you can attach the cgroups to "ioband1" and assign them weights.

```
# dmsetup message ioband1 0 type cgroup
# dmsetup message ioband1 0 attach 1
# dmsetup message ioband1 0 attach 2
# dmsetup message ioband1 0 weight 1:30
# dmsetup message ioband1 0 weight 2:60
```

You can find the manual of dm-ioband at
http://people.valinux.co.jp/~ryov/dm-ioband/manual/index.html.
But the user interface for the bio cgroup is temporal and it will be
changed after the io_context support.

Thank you,
Hirokazu Takahashi.