Subject: Re: [PATCH 2/4] Block I/O tracking
Posted by Hirokazu Takahashi on Tue, 18 Mar 2008 12:15:05 GMT
View Forum Message <> Reply to Message

Hi,

> > > And, blist seems to be just used for force_empty.
> > > Do you really need this ? no alternative ?
> >
> > I selected this approach because it was the simplest way for the
> > first implementation.
> >
> > I've been also thinking about what you pointed.
> > If you don't mind taking a long time to remove a bio cgroup, it will be
> > the easiest way that you can scan all pages to find the pages which
> > belong to the cgroup and delete them. It may be enough since you may
> > say it will rarely happen. But it might cause some trouble on machines
> > with huge memory.
> >
> Hmm, force_empty itself is necessary ?

It is called when bio cgroups are removed.
With the current implementation, when you delete a bio cgroup,
the bio_cgroup members of page_cgroups which point the cgroup
have to be cleared.

So I'm looking for another way like:
- Use some kind of id instead of a pointer to a bio cgroup,
  so you can check whether the id is valid before you use it.
- Don't free the bio cgroup until all the pages referring to
  the cgroup.

I also want to implement that if you find a page whose cgroup is
already removed, the page should be assigned to a new cgroup.

Thank you,
Hirokazu Takahashi.