

---

Subject: Re: [PATCH 3/5] page\_cgroup: make page tracking available for blkio  
Posted by [Andrea Righi](#) on Tue, 22 Feb 2011 23:08:40 GMT

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

---

On Tue, Feb 22, 2011 at 04:22:53PM -0500, Vivek Goyal wrote:

> On Tue, Feb 22, 2011 at 06:12:54PM +0100, Andrea Righi wrote:

> > The page\_cgroup infrastructure, currently available only for the memory  
> > cgroup controller, can be used to store the owner of each page and  
> > opportunely track the writeback IO. This information is encoded in  
> > the upper 16-bits of the page\_cgroup->flags.

> >

> > A owner can be identified using a generic ID number and the following  
> > interfaces are provided to store a retrieve this information:

> >

> > unsigned long page\_cgroup\_get\_owner(struct page \*page);  
> > int page\_cgroup\_set\_owner(struct page \*page, unsigned long id);  
> > int page\_cgroup\_copy\_owner(struct page \*npage, struct page \*opage);

> >

> > The blkio.throttle controller can use the cgroup css\_id() as the owner's  
> > ID number.

> >

> > Signed-off-by: Andrea Righi <[arighi@develer.com](mailto:arighi@develer.com)>

> > ---

> > block/Kconfig | 2 +  
> > block/blk-cgroup.c | 6 ++  
> > include/linux/memcontrol.h | 6 ++  
> > include/linux/mmzone.h | 4 +-  
> > include/linux/page\_cgroup.h | 33 ++++++++-  
> > init/Kconfig | 4 +  
> > mm/Makefile | 3 +-  
> > mm/memcontrol.c | 6 ++  
> > mm/page\_cgroup.c | 129 ++++++-----  
> > 9 files changed, 176 insertions(+), 17 deletions(-)

> >

> > diff --git a/block/Kconfig b/block/Kconfig

> > index 60be1e0..1351ea8 100644

> > --- a/block/Kconfig

> > +++ b/block/Kconfig

> > @@ -80,6 +80,8 @@ config BLK\_DEV\_INTEGRITY

> > config BLK\_DEV\_THROTTLING

> > bool "Block layer bio throttling support"

> > depends on BLK\_CGROUP=y && EXPERIMENTAL

> > + select MM\_OWNER

> > + select PAGE\_TRACKING

> > default n

> > ---help---

> > Block layer bio throttling support. It can be used to limit

> > diff --git a/block/blk-cgroup.c b/block/blk-cgroup.c

```
> > index f283ae1..5c57f0a 100644
> > --- a/block/blk-cgroup.c
> > +++ b/block/blk-cgroup.c
> > @@ -107,6 +107,12 @@ blkio_policy_search_node(const struct blkio_cgroup *blkcg, dev_t
dev,
> >     return NULL;
> > }
> >
> > +bool blkio_cgroup_disabled(void)
> > +{
> > +    return blkio_subsys.disabled ? true : false;
> > +}
> > +EXPORT_SYMBOL_GPL(blkio_cgroup_disabled);
> > +
>
> I think there should be option to just disable this asyn feature of
> blkio controller. So those who don't want it (running VMs with cache=none
> option) and don't want to take the memory reservation hit should be
> able to disable just ASYNC facility of blkio controller and not
> the whole blkio controller facility.
```

Definitely a better choice.

OK, I'll apply all your suggestions and post a new version of the patch.

Thanks for the review!

-Andrea

---

Containers mailing list  
Containers@lists.linux-foundation.org  
<https://lists.linux-foundation.org/mailman/listinfo/containers>

---