
Subject: Re: [PATCH 3/5] page_cgroup: make page tracking available for blkio
Posted by [Vivek Goyal](#) on Tue, 22 Feb 2011 21:22:53 GMT
[View Forum Message](#) <> [Reply to Message](#)

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
> opportunistically 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 and 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>

> ---

> 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.

Thanks
Vivek

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