Subject: Re: [PATCH 3/5] page_cgroup: make page tracking available for blkio
Posted by KAMEZAWA Hiroyuki on Wed, 23 Feb 2011 23:58:05 GMT
View Forum Message <> Reply to Message

On Wed, 23 Feb 2011 09:59:11 +0100
Andrea Righi <arighi@develer.com> wrote:

> >
> > I wonder I can make pc->mem_cgroup to be pc->memid(16bit), then,
> > ==
> > static inline struct mem_cgroup *get_memcg_from_pc(struct page_cgroup *pc)
> > {
> >  struct cgroup_subsys_state *css = css_lookup(&mem_cgroup_subsys, pc->memid);
> >  return container_of(css, struct mem_cgroup, css);
> > }
> > ==
> > Overhead will be seen at updating file statistics and LRU management.
> >
> > But, hmm, can't you do that tracking without page_cgroup ?
> > Because the number of dirty/writeback pages are far smaller than total pages,
> > chasing I/O with dynamic structure is not very bad..
> >
> > prepareing [pfn -> blkio] record table and move that information to struct bio
> > in dynamic way is very difficult ?
>
> This would be ok for dirty pages, but consider that we're also tracking
> anonymous pages. So, if we want to control the swap IO we actually need
> to save this information for a lot of pages and at the end I think we'll
> basically duplicate the page_cgroup code.
>

swap io is always started with bio and the task/mm_struct.
So, if we can record information in bio, no page tracking is required.
You can record information to bio just by reading mm->owner.

Thanks,
-Kame

_____
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containe rs