Subject: Re: [PATCH v2 02/11] memcg: Reclaim when more than one page needed.
Posted by Michal Hocko on Mon, 13 Aug 2012 13:10:14 GMT
View Forum Message <> Reply to Message

On Mon 13-08-12 12:05:38, Glauber Costa wrote:
> On 08/10/2012 10:54 PM, Michal Hocko wrote:
> > On Thu 09-08-12 17:01:10, Glauber Costa wrote:
> >> From: Suleiman Souhlal <ssouhlal@FreeBSD.org>
> >>
> >> mem_cgroup_do_charge() was written before kmem accounting, and expects
> >> three cases: being called for 1 page, being called for a stock of 32
> >> pages, or being called for a hugepage. If we call for 2 or 3 pages (and
> >> both the stack and several slabs used in process creation are such, at
> >> least with the debug options I had), it assumed it's being called for
> >> stock and just retried without reclaiming.
> >>
> >> Fix that by passing down a minsize argument in addition to the csize.
> >>
> >> And what to do about that (csize == PAGE_SIZE && ret) retry? If it's
> >> needed at all (and presumably is since it's there, perhaps to handle
> >> races), then it should be extended to more than PAGE_SIZE, yet how far?
> >> And should there be a retry count limit, of what? For now retry up to
> >> COSTLY_ORDER (as page_alloc.c does) and make sure not to do it if
> >> __GFP_NORETRY.
> >>
> >> [v4: fixed nr pages calculation pointed out by Christoph Lameter]
> >>
> >> Signed-off-by: Suleiman Souhlal <suleiman@google.com>
> >> Signed-off-by: Glauber Costa <glommer@parallels.com>
> >> Reviewed-by: Kamezawa Hiroyuki <kamezawa.hiroyu@jp.fujitsu.com>
> >
> > I am not happy with the min_pages argument but we can do something more
> > clever later.
> >
> > Acked-by: Michal Hocko <mhocko@suse.cz>
> >
>
> I am a bit confused here. Does your ack come before or after your other
> comments on this patch?

Heh, it was hard Friday ;) Yes, it was after the mind fart...
--
Michal Hocko
SUSE Labs