Subject: Re: [PATCH v2 02/11] memcg: Reclaim when more than one page needed.
Posted by KAMEZAWA Hiroyuki on Fri, 10 Aug 2012 16:49:25 GMT
View Forum Message <> Reply to Message

(2012/08/11 0:42), Michal Hocko wrote:
> On Thu 09-08-12 17:01:10, Glauber Costa wrote:
> [...]
>> @@ -2317,18 +2318,18 @@ static int mem_cgroup_do_charge(struct mem_cgroup *memcg, gfp_t gfp_mask,
>>   } else
>>   mem_over_limit = mem_cgroup_from_res_counter(fail_res, res);
>>   /*
>> -   * nr_pages can be either a huge page (HPAGE_PMD_NR), a batch
>> -   * of regular pages (CHARGE_BATCH), or a single regular page (1).
>> -   *
>>   * Never reclaim on behalf of optional batching, retry with a
>>   * single page instead.
>>   */
>> - if (nr_pages == CHARGE_BATCH)
>> + if (nr_pages > min_pages)
>>   return CHARGE_RETRY;
>
> This is dangerous because THP charges will be retried now while they
> previously failed with CHARGE_NOMEM which means that we will keep
> attempting potentially endlessly.

with THP, I thought nr_pages == min_pages, and no retry.

> Why cannot we simply do if (nr_pages < CHARGE_BATCH) and get rid of the
> min_pages altogether?

Hm, I think a slab can be larger than CHARGE_BATCH.

> Also the comment doesn't seem to be valid anymore.
>
I agree it's not clean. Because our assumption on nr_pages are changed,
I think this behavior should not depend on nr_pages value..
Shouldn't we have a flag to indicate "trial-for-batched charge" ?

Thanks,
-Kame