Subject: Re: [PATCH 10/11] memcg: allow a memcg with kmem charges to be destructed.
Posted by Glauber Costa on Mon, 25 Jun 2012 22:25:46 GMT

On 06/25/2012 10:34 PM, Tejun Heo wrote:
> On Mon, Jun 25, 2012 at 06:15:27PM +0400, Glauber Costa wrote:
>> Because the ultimate goal of the kmem tracking in memcg is to
>> track slab pages as well, we can't guarantee that we'll always
>> be able to point a page to a particular process, and migrate
>> the charges along with it - since in the common case, a page
>> will contain data belonging to multiple processes.
>>
>> Because of that, when we destroy a memcg, we only make sure
>> the destruction will succeed by discounting the kmem charges
>> from the user charges when we try to empty the cgroup.
>>
>> Signed-off-by: Glauber Costa <glommer@parallels.com>
>> CC: Christoph Lameter <cl@linux.com>
>> CC: Pekka Enberg <penberg@cs.helsinki.fi>
>> CC: Michal Hocko <mhocko@suse.cz>
>> CC: Kamezawa Hiroyuki <kamezawa.hiroyu@jp.fujitsu.com>
>> CC: Johannes Weiner <hannes@cmpxchg.org>
>> CC: Suleiman Souhlal <suleiman@google.com>
>> ---
>>   mm/memcontrol.c | 10 +++++++++-
>>   1 file changed, 9 insertions(+), 1 deletion(-)
>>
>> diff --git a/mm/memcontrol.c b/mm/memcontrol.c
>> index a6a440b..bb9b6fe 100644
>> --- a/mm/memcontrol.c
>> +++ b/mm/memcontrol.c
>> @@ -598,6 +598,11 @@ static void disarm_kmem_keys(struct mem_cgroup *memcg)
>>   {
>>   if (test_bit(KMEM_ACCOUNTED_THIS, &memcg->kmem_accounted))
>>   static_key_slow_dec(&mem_cgroup_kmem_enabled_key);
>> + /*
>> + * This check can't live in kmem destruction function,
>> + * since the charges will outlive the cgroup
>> + */
>> + BUG_ON(res_counter_read_u64(&memcg->kmem, RES_USAGE) != 0);
>
> WARN_ON() please. Misaccounted kernel usually is better than dead
> kernel.
>

You're absolutely right, will change.