
Subject: Re: [PATCH 07/10] memcg: Stop res_counter underflows.

Posted by [Suleiman Souhlal](#) on Tue, 28 Feb 2012 23:07:39 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Tue, Feb 28, 2012 at 5:31 AM, Glauber Costa <glommer@parallels.com> wrote:

> I don't fully understand this.

> To me, the whole purpose of having a cache tied to a memcg, is that we know

> all allocations from that particular cache should be billed to a specific

> memcg. So after a cache is created, and has an assigned memcg,

> what's the point in bypassing it to root?

>

> It smells like you're just using this to circumvent something...

In the vast majority of the cases, we will be able to account to the cgroup.

However, there are cases when `__mem_cgroup_try_charge()` is not able to do so, like when the task is being killed.

When this happens, the allocation will not get accounted to the cgroup, but the slab accounting code will still think the page belongs to the memcg's `kmem_cache`.

So, when we go to free the page, we assume that the page belongs to the memcg and uncharge it, even though it was never charged to us in the first place.

This is the situation this patch is trying to address, by keeping a counter of how much memory has been bypassed like this, and uncharging from the root if we have any outstanding bypassed memory.

Does that make sense?

-- Suleiman
