
Subject: [RFC 4/4] memcg: do not walk all the way to the root for memcg
Posted by [Glauber Costa](#) on Tue, 25 Sep 2012 08:52:53 GMT

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

Since the root is special anyway, and we always get its figures from global counters anyway, there is no make all cgroups its descendants, wrt res_counters. The sad effect of doing that is that we need to lock the root for all allocations, since it is a common ancestor of everybody.

Not having the root as a common ancestor should lead to better scalability for not-uncommon case of tasks in the cgroup being node-bound to different nodes in NUMA systems.

Signed-off-by: Glauber Costa <glommer@parallels.com>
CC: Michal Hocko <mhocko@suse.cz>
CC: Kamezawa Hiroyuki <kamezawa.hiroyu@jp.fujitsu.com>
CC: Johannes Weiner <hannes@cmpxchg.org>
CC: Mel Gorman <mgorman@suse.de>
CC: Andrew Morton <akpm@linux-foundation.org>

mm/memcontrol.c | 2 +-
1 file changed, 1 insertion(+), 1 deletion(-)

diff --git a/mm/memcontrol.c b/mm/memcontrol.c
index f8115f0..829ea9e 100644

--- a/mm/memcontrol.c

+++ b/mm/memcontrol.c

```
@@ -5720,7 +5720,7 @@ mem_cgroup_create(struct cgroup *cont)
    static_key_slow_inc(&memcg_in_use_key);
}
```

```
- if (parent && parent->use_hierarchy) {
```

```
+ if (parent && !mem_cgroup_is_root(parent) && parent->use_hierarchy) {
    struct mem_cgroup __maybe_unused *p;
```

```
    res_counter_init(&memcg->res, &parent->res);
```

```
--
```

1.7.11.4
