
Subject: Re: [PATCH 6/6 mm] memcggroup: revert swap_state mods
Posted by [Hugh Dickins](#) on Mon, 12 Nov 2007 04:57:03 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Fri, 9 Nov 2007, KAMEZAWA Hiroyuki wrote:
> On Fri, 9 Nov 2007 07:14:22 +0000 (GMT)
> Hugh Dickins <hugh@veritas.com> wrote:
>
> > If we're charging rss and we're charging cache, it seems obvious that
> > we should be charging swapcache - as has been done. But in practice
> > that doesn't work out so well: both swapin readahead and swapoff leave
> > the majority of pages charged to the wrong cgroup (the cgroup that
> > happened to read them in, rather than the cgroup to which they belong).
>
> Thank you. I welcome this patch :)

Thank you! But perhaps less welcome if I don't confirm...

> Could I confirm a change in the logic ?
>
> * Before this patch, wrong swapcache charge is added to one who
> called `try_to_free_page()`.

`try_to_free_pages`? No, I don't think any wrong charge was made there. It was when reading in swap pages. The usual way is by `swapin_readahead`, which reads in a cluster of swap pages, which are quite likely to belong to different memcgroups, but were all charged to the one which is doing the fault on its target page. Another way is in `swapoff`, where they all got charged to whoever was doing the `swapoff` (and the charging in `unuse_pte` was a no-op).

> * After this patch, anonymous page's charge will drop to 0 when
> `page_remove_rmap()` is called.

Yes, when its final (usually its only) `page_remove_rmap` is called.

Hugh

Containers mailing list
Containers@lists.linux-foundation.org
<https://lists.linux-foundation.org/mailman/listinfo/containers>
