Subject: Re: [RFC] [-mm PATCH] Memory controller fix swap charging context in unuse_pte()
Posted by Balbir Singh on Wed, 24 Oct 2007 12:14:42 GMT
View Forum Message <> Reply to Message

Hugh Dickins wrote:
> On Mon, 15 Oct 2007, Balbir Singh wrote:
>> Hugh Dickins wrote:
>>> --- 2.6.23-rc8-mm2/mm/swapfile.c 2007-09-27 12:03:36.000000000 +0100
>>> +++ linux/mm/swapfile.c 2007-10-07 14:33:05.000000000 +0100
>>> @@ -507,11 +507,23 @@ unsigned int count_swap_pages(int type,
>>> * just let do_wp_page work it out if a write is requested later - to
>>> * force COW, vm_page_prot omits write permission from any private vma.
>>> */
>>> -static int unuse_pte(struct vm_area_struct *vma, pte_t *pte,
>>> +static int unuse_pte(struct vm_area_struct *vma, pmd_t *pmd,
>>> unsigned long addr, swp_entry_t entry, struct page *page)
> ...
>> I tested this patch and it seems to be working fine. I tried swapoff -a
>> in the middle of tests consuming swap. Not 100% rigorous, but a good
>> test nevertheless.
>>
>> Tested-by: Balbir Singh <balbir@linux.vnet.ibm.com>
>
> Thanks, Balbir. Sorry for the delay. I've not forgotten our
> agreement that I should be splitting it into before-and-after
> mem cgroup patches. But it's low priority for me until we're
> genuinely assigning to a cgroup there. Hope to get back to
> looking into that tomorrow, but no promises.
>

No Problem. We have some time with this one.

> I think you still see no problem, where I claim that simply
> omitting the mem charge mods from mm/swap_state.c leads to OOMs?
> Maybe our difference is because my memhog in the cgroup is using
> more memory than RAM, not just more memory than allowed to the
> cgroup. I suspect that arrives at a state (when the swapcache
> pages are not charged) where it cannot locate the pages it needs
> to reclaim to stay within its limit.
>

Yes, in my case there I use memory less than RAM and more than that
is allowed by the cgroup. It's quite possible that in your case the
swapcache has grown significantly without any limit/control on it.
The memhog program is using memory at a rate much higher than the
rate of reclaim. Could you share your memhog program, please?
In the use case you've mentioned/tested, having these mods to

control swapcache is actually useful, right?

Could you share your major objections at this point with the memory
controller at this point. I hope to be able to look into/resolve them
as my first priority in my list of items to work on.

> Hugh

--
Warm Regards,
Balbir Singh
Linux Technology Center
IBM, ISTL

_____

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