

---

Subject: Re: [RFC] [-mm PATCH] Memory controller fix swap charging context in  
unuse\_pte()

Posted by [Balbir Singh](#) on Tue, 30 Oct 2007 18:28:20 GMT

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

---

Hugh Dickins wrote:

> On Tue, 30 Oct 2007, Balbir Singh wrote:

>> At this momemnt, I suspect one of two things

>>

>> 1. Our mods to swap\_state.c are different

>

> I believe they're the same (just take swap\_state.c back to how it  
> was without mem\_cgroup mods) - or would be, if after finding this  
> effect I hadn't added a "swap\_in\_cg" switch to move between the  
> two behaviours to study it better (though I do need to remember  
> to swapoff and swapon between the two: sometimes I do forget).

>

>> 2. Our configuration is different, main-memory to swap-size ratio

>

> I doubt the swappiness is relevant: just so long as there's some (a  
> little more than 200M I guess); I've got 1GB-2GB on different boxes.

>

I agree, just wanted to make sure that there is enough swap

> There may well be something about our configs that's significantly  
> different. I'd failed to mention SMP (4 cpu), and that I happen  
> to have /proc/sys/vm/swappiness 100; but find it happens on UP  
> also, and when I go back to default swappiness 60.

>

OK.. so those are out of the equation

> I've reordered your mail for more dramatic effect...

>> On a real box - a powerpc machine that I have access to

>

> I've tried on 3 Intel and 1 PowerPC now: the Intels show the OOMs  
> and the PowerPC does not. I rather doubt it's an Intel versus  
> PowerPC issue as such, but interesting that we see the same.

>

Very surprising, I am surprised that it's architecture dependent.

Let me try and grab an Intel box and try.

>> 1. I don't see the OOM with the mods removed (I have swap

>> space at-least twice of RAM - with mem=512M, I have at-least

>> 1G of swap).

>

> mem=512M with 1G of swap, yes, I'm the same.  
>  
>> 2. Running under the container is much much faster than running  
>> swapout in the root container. The machine is almost unusable  
>> if swapout is run under the root container  
>  
> That's rather interesting, isn't it? Probably irrelevant to the  
> OOM issue we're investigating, but worthy of investigation in itself.  
>

Yes, it irrelevant, but I find it to be a good use case for using the memory controller :-). I found that kswapd running at prio -5, seemed to hog quite a bit of the CPU. But it needs more independent investigation, like you've suggested.

> Maybe I saw the same on the PowerPC: I simply forgot to set up the  
> cgroup one time, and my sequence of three swapouts (sometimes only  
> two out of three OOM, on those boxes that do OOM) seemed to take a  
> very long time (but I wasn't trying to do anything else on it at  
> the same time, so didn't notice if it was "unusable").  
>  
> I'll probe on.  
>

Me too.. I'll try and acquire a good x86\_64 box and test on it.

> Hugh  
>  
> --  
> To unsubscribe, send a message with 'unsubscribe linux-mm' in  
> the body to majordomo@kvack.org. For more info on Linux MM,  
> see: <http://www.linux-mm.org/> .  
> Don't email: [dont@kvack.org](mailto:dont@kvack.org) email@kvack.org

--

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>

---