Subject: Re: [RFC][PATCH] another swap controller for cgroup
Posted by Daisuke Nishimura on Tue, 25 Mar 2008 06:46:28 GMT
View Forum Message <> Reply to Message

Balbir Singh wrote:
> Daisuke Nishimura wrote:
>> Daisuke Nishimura wrote:
>>> Hi, Yamamoto-san.
>>>
>>> I'm reviewing and testing your patch now.
>>>
>> In building kernel infinitely(in a cgroup of
>> memory.limit=64M and swap.limit=128M, with swappiness=100),
>> almost all of the swap (1GB) is consumed as swap cache
>> after a day or so.
>> As a result, processes are occasionally OOM-killed even when
>> the swap.usage of the group doesn't exceed the limit.
>>
>> I don't know why the swap cache uses up swap space.
>> I will test whether a similar issue happens without your patch.
>> Do you have any thoughts?
>>
>>
>> BTW, I think that it would be better, in the sence of
>> isolating memory resource, if there is a framework
>> to limit the usage of swap cache.
>
> We had this earlier, but dropped it later due to issues related to swap
> readahead and assigning the pages to the correct cgroup.
>
Yes, I know.

In my swap subsystem posted before, I charge swap entries
and remember to which cgroup each swap entries is charged
in an array of pointers. So swap caches is charged as swap
not memory, and swap usage including swap cache can be accounted.

There may be better solution, and one of the issue of
my implementation is swap_cgroup_chage() returns error
before reclaiming swap entries which are only used
by swap caches.
I'm considering this issue now.

Thanks,
Daisuke Nishimura.

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