Subject: Re: [PATCH 0/4] swapcgroup(v2)
Posted by Balbir Singh on Fri, 23 May 2008 03:59:23 GMT
View Forum Message <> Reply to Message

KAMEZAWA Hiroyuki wrote:
> On Thu, 22 May 2008 22:26:55 -0400
> Rik van Riel <riel@redhat.com> wrote:
>
>> Even worse is that a cgroup has NO CONTROL over how much
>> of its memory is kept in RAM and how much is swapped out.

We used to have a control on the swap cache pages as well, but their
implementation needed more thought

> Could you explain "NO CONTROL" ? cgroup has LRU....
> 'how mucch memory should be swapped out from memory' is well controlled
> in the VM besides LRU logic ?
>
>> This kind of decision is made on a system-wide basis by
>> the kernel, dependent on what other processes in the system
>> are doing. There also is no easy way for a cgroup to reduce
>> its swap use, unlike with other resources.
>>

One option is to limit the virtual address space usage of the cgroup to ensure
that swap usage of a cgroup will *not* exceed the specified limit. Along with a
good swap controller, it should provide good control over the cgroup's memory usage.

>
>> In what scenario would you use a resource controller that
>> rewards a group for reaching its limit?
>>
>> How can the cgroup swap space controller help sysadmins
>> achieve performance or fairness goals on a system?
>>
> Perforamnce is not the first goal of this swap controller, I think.
> This is for resouce isolation/overcommiting.
>
> 1. Some _crazy_ people considers swap as very-slow-memory resource ;)
>   I don't think so but I know there are tons of people....
>
> 2. Resource Isolation.
>   When a cgroup has memory limitation, it can create tons of swap.
>   For example, limit a cgroup's memory to be 128M and malloc 3G bytes.
>   2.8Gbytes of swap will be used _easily_. A process can use up all swap.
>   In that case, other process can't use swap.
>
> IIRC, a man shown his motivation to controll swap in OLS2007/BOF as following.

> ==
> Consider following system. (and there is no swap controller.)
> Memory 4G. Swap 1G. with 2 cgroups A, B.
>
> state 1) swap is not used.
> A....memory limit to be 1G no swap usage memory_usage=0M
> B....memory limit to be 1G no swap usage memory_usage=0M
>
> state 2) Run a big program on A.
> A....memory limit to be 1G and try to use 1.7G. uses 700MBytes of swap.
> memory_usage=1G swap_usage=700M
> B....memory_usage=0M
>
> state 3) A some of programs ends in 'A'
> A....memory_usage=500M swap_usage=700M
> B....memory_usage=0M.
>
> state 4) Run a big program on B.
> A...memory_usage=500M swap_usage=700M.
> B...memory_usage=1G swap_usage=300M
>
> Group B can only use 1.3G because of unfair swap use of group A.
> But users think why A uses 700M of swap with 500M of free memory....
>
> If we don't have limitation to swap, we'll have to innovate a way to move swap
> to memory in some reasonable logic.
>
> Thanks,
> -Kame

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

_____