Subject: Re: [PATCH 2/2] Make res_counter hierarchical
Posted by Paul Menage on Tue, 11 Mar 2008 08:57:43 GMT
View Forum Message <> Reply to Message

On Tue, Mar 11, 2008 at 1:15 AM, Pavel Emelyanov <xemul@openvz.org> wrote:
>
> <mem_couter_0>
> + -- <swap_counter_0>
> + -- <mem_counter_1>
> | + -- <swap_counter_1>
> | + -- <mem_counter_11>
> | | + -- <swap_counter_11>
> | + -- <mem_counter_12>
> | + -- <swap_counter_12>
> + -- <mem_counter_2>
> | + -- <swap_counter_2>
> | + -- <mem_counter_21>
> | | + -- <swap_counter_21>
> | + -- <mem_counter_22>
> | + -- <swap_counter_22>
> + -- <mem_counter_N>
> + -- <swap_counter_N>
> + -- <mem_counter_N1>
> | + -- <swap_counter_N1>
> + -- <mem_counter_N2>
> + -- <swap_counter_N2>
>

The idea of hierarchy is good, but I don't think this particular hierarchy works for memory.

Main memory and swap space are very different resources, with very different performance characteristics. Suppose you have a 2G machine, and you want to guarantee each job 1GB of main memory, plus give them the option of 1GB of swap for when they go over the 1G main memory limit. With the hierarchy given above, you've need to give each job a 2GB mem.limit and a 1GB swap.limit, and so there would be no main memory isolation.

My feeling is that people are going to want to limit swap and main memory usage as two independent resource hierarchies more often than they're going to want to limit overall virtual memory. But assuming that there are people who need to do the latter, then you should make it configurable how the hierarchies fit together.

Alternatively, you could make it possible for a res_counter to have multiple parents (each of which constrains the overall usage of it and its siblings), and have three counters for each cgroup:

- vm_counter: overall virtual memory limit for group, parent = parent_mem_cgroup->vm_counter

- mem_counter: main memory limit for group, parents = vm_counter, parent_mem_cgroup->mem_counter

- swap_counter: swap limit for group, parents = vm_counter, parent_mem_cgroup->swap_counter

Paul

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