Subject: Re: [PATCH] Simplify memory controller and resource counter I/O
Posted by Balbir Singh on Fri, 05 Oct 2007 04:04:23 GMT
View Forum Message <> Reply to Message

Paul Menage wrote:
> On 10/4/07, Balbir Singh <balbir@linux.vnet.ibm.com> wrote:
>> Paul Menage wrote:
>>> On 10/4/07, Balbir Singh <balbir@linux.vnet.ibm.com> wrote:
>>>> Forbidding writing to the root resource counter is a policy decision
>>>> I am unable to make up my mind about. It sounds right, but unless
>>>> we have a notion of unlimited resources, I am a bit concerned about
>>>> taking away this flexibility.
>>> One big reason for doing this is to make virtualization easier - if
>>> you expect not to be able to write to your root cgroup's limits files,
>>> then it's easier to make them non-writeable for a virtual server.
>>>
>> Can't we handle that through file system permissions? virtual servers
>> will not run as root
>
> They'll probably run as root in their own user namespace if at all.
> But that's the point - if userspace in general expects root cgroup
> limits to not be writeable (the same way that root cpusets
> cpus/mems_allowed files aren't writeable) then virtual servers will
> break less.
>

In that case, let's have a value that says RES_COUNTER_INFINITY
and set the root to that value and make the root cgroup limits
read-only.

>> But system administrators deal with memory in MB and GB. When you go
>> to buy memory, you don't specify, I need 1 << 30 or 2^30 bytes of
>> memory :-). Most administrators track their memory using these
>> quantifiers.
>
> OK, so maybe we should just fold a call to memparse() into
> cgroup_write_uint? Then we could use the plain write_uint() method in
> the control file?
>

Yes, either that way or add a strategy function, that would take
the string input from the user and convert it to unsigned long long
value. I am ok with either approach.

>>>> Do read_uint() and write_uint(), just read and write unsigned
>>>> integers?
>>> Correct.
>>>

>> Oops.. that would be problem, what if I wanted to set my limit to
>> unsigned long long max?
>
> Sorry, I wasn't getting your point about the sizing. No, they're u64
> values. (And I guess could be changed to unsigned long long if people
> preferred).
>

I would prefer unsigned long long, but we could get more opinions.

> Paul

--
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