
Subject: Re: [PATCH 08/10] memcg: Add
CONFIG_CGROUP_MEM_RES_CTLR_KMEM_ACCT_ROOT.
Posted by [Glauber Costa](#) on Wed, 29 Feb 2012 17:09:25 GMT
[View Forum Message](#) <> [Reply to Message](#)

On 02/28/2012 08:36 PM, Suleiman Souhlal wrote:
> On Tue, Feb 28, 2012 at 5:34 AM, Glauber Costa<glommer@parallels.com> wrote:
>> On 02/27/2012 07:58 PM, Suleiman Souhlal wrote:
>>>
>>> This config option dictates whether or not kernel memory in the
>>> root cgroup should be accounted.
>>>
>>> This may be useful in an environment where everything is supposed to be
>>> in a cgroup and accounted for. Large amounts of kernel memory in the
>>> root cgroup would indicate problems with memory isolation or accounting.
>>
>>
>> I don't like accounting this stuff to the root memory cgroup. This causes
>> overhead for everybody, including people who couldn't care less about memcg.
>>
>> If it were up to me, we would simply not account it, and end of story.
>>
>> However, if this is terribly important for you, I think you need to at
>> least make it possible to enable it at runtime, and default it to disabled.
>
> Yes, that is why I made it a config option. If the config option is
> disabled, that memory does not get accounted at all.

Doesn't work. In reality, most of the distributions enable those stuff
if there is the possibility that someone will end up using. So everybody
gets to pay the penalty.

> Making it configurable at runtime is not ideal, because we would
> prefer slab memory that was allocated before cgroups are created to
> still be counted toward root.
>

Again: Why is that you really need it ? Accounting slab to the root
cgroup feels quite weird to me

Subject: Re: [PATCH 08/10] memcg: Add
CONFIG_CGROUP_MEM_RES_CTLR_KMEM_ACCT_ROOT.
Posted by [Suleiman Souhlal](#) on Wed, 29 Feb 2012 19:24:05 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Wed, Feb 29, 2012 at 9:09 AM, Glauber Costa <glommer@parallels.com> wrote:
> On 02/28/2012 08:36 PM, Suleiman Souhlal wrote:

```

>>
>> On Tue, Feb 28, 2012 at 5:34 AM, Glauber Costa<glommer@parallels.com>
>> wrote:
>>>
>>> On 02/27/2012 07:58 PM, Suleiman Souhlal wrote:
>>>>
>>>>
>>>> This config option dictates whether or not kernel memory in the
>>>> root cgroup should be accounted.
>>>>
>>>> This may be useful in an environment where everything is supposed to be
>>>> in a cgroup and accounted for. Large amounts of kernel memory in the
>>>> root cgroup would indicate problems with memory isolation or accounting.
>>>
>>>
>>>
>>> I don't like accounting this stuff to the root memory cgroup. This causes
>>> overhead for everybody, including people who couldn't care less about
>>> memcg.
>>>
>>> If it were up to me, we would simply not account it, and end of story.
>>>
>>> However, if this is terribly important for you, I think you need to at
>>> least make it possible to enable it at runtime, and default it to
>>> disabled.
>>
>>
>> Yes, that is why I made it a config option. If the config option is
>> disabled, that memory does not get accounted at all.
>
>
> Doesn't work. In reality, most of the distributions enable those stuff if
> there is the possibility that someone will end up using. So everybody gets
> to pay the penalty.
>
>
>> Making it configurable at runtime is not ideal, because we would
>> prefer slab memory that was allocated before cgroups are created to
>> still be counted toward root.
>>
>
> Again: Why is that you really need it ? Accounting slab to the root cgroup
> feels quite weird to me

```

Because, for us, having large amounts of unaccounted memory is a "bug", and we would like to know when it happens.
Also, we want to know how much memory is actually available in the machine for jobs (sum of(accounted memory in containers) - unaccounted

kernel memory).

That said, I will drop this patch from the series for now.

-- Suleiman
