Subject: Re: [PATCH v2 04/11] kmem accounting basic infrastructure
Posted by Glauber Costa on Mon, 13 Aug 2012 08:36:51 GMT
View Forum Message <> Reply to Message

On 08/10/2012 09:02 PM, Kamezawa Hiroyuki wrote:
> (2012/08/09 22:01), Glauber Costa wrote:
>> This patch adds the basic infrastructure for the accounting of the slab
>> caches. To control that, the following files are created:
>>
>> * memory.kmem.usage_in_bytes
>> * memory.kmem.limit_in_bytes
>> * memory.kmem.failcnt
>> * memory.kmem.max_usage_in_bytes
>>
>> They have the same meaning of their user memory counterparts. They
>> reflect the state of the "kmem" res_counter.
>>
>> The code is not enabled until a limit is set. This can be tested by the
>> flag "kmem_accounted". This means that after the patch is applied, no
>> behavioral changes exists for whoever is still using memcg to control
>> their memory usage.
>>
>> We always account to both user and kernel resource_counters. This
>> effectively means that an independent kernel limit is in place when the
>> limit is set to a lower value than the user memory. A equal or higher
>> value means that the user limit will always hit first, meaning that kmem
>> is effectively unlimited.
>>
>> People who want to track kernel memory but not limit it, can set this
>> limit to a very high number (like RESOURCE_MAX - 1page - that no one
>> will ever hit, or equal to the user memory)
>>
>> Signed-off-by: Glauber Costa <glommer@parallels.com>
>> CC: Michal Hocko <mhocko@suse.cz>
>> CC: Johannes Weiner <hannes@cmpxchg.org>
>> Reviewed-by: Kamezawa Hiroyuki <kamezawa.hiroyu@jp.fujitsu.com>
>
> Could you add a patch for documentation of this new interface and a text
> explaining the behavior of "kmem_accounting" ?
>
> Hm, my concern is the difference of behavior between user page accounting and
> kmem accounting...but this is how tcp-accounting is working.
>
> Once you add Documentation, it's okay to add my Ack.
>
I plan to add documentation in a separate patch. Due to that, can I add
your ack to this patch here?

Also, I find that the description text in patch0 grew to be quite informative and complete. I plan to add that to the documentation if that is ok with you