Subject: Re: [PATCH 0/4] Kernel memory accounting container (v3)
Posted by Pavel Emelianov on Tue, 18 Sep 2007 06:56:21 GMT

Christoph Lameter wrote:
> On Tue, 18 Sep 2007, Balbir Singh wrote:
>
>> I've wondered the same thing and asked the question. Pavel wrote
>> back to me saying
>>
>> "The pages that are full of objects are not linked in any list
>> in kmem_cache so we just cannot find them."
>
> That is true for any types of slab cache and not restricted to kmalloc
> slabs. SLUB can be switched into a mode where it provides these lists
> (again at a performance penalty).
>
> But I thought we generate the counters at alloc and free time? So why do
> we need to traverse the object lists?

When we make echo 0 > /sys/slab/xxx/cache_notify we want all the objects
to be unaccounted back immediately. Even __free_slab() won't catch this
because the SLAB_NOTIFY flag will be turned off for this cache. So we
have to walk all the objects and unaccount them.