Subject: Re: [PATCH v2 04/29] slub: always get the cache from its page in kfree
Posted by Glauber Costa on Fri, 11 May 2012 19:11:05 GMT
View Forum Message <> Reply to Message

On 05/11/2012 04:09 PM, Christoph Lameter wrote:
> On Fri, 11 May 2012, Glauber Costa wrote:
>
>> On 05/11/2012 03:56 PM, Christoph Lameter wrote:
>>> On Fri, 11 May 2012, Glauber Costa wrote:
>>>
>>>> So we don't mix pages from multiple memcgs in the same cache - we believe
>>>> that
>>>> would be too confusing.
>>>
>>> Well subsystem create caches and other things that are shared between
>>> multiple processes. How can you track that?
>>
>> Each process that belongs to a memcg triggers the creation of a new child kmem
>> cache.
>
> I see that. But there are other subsystems from slab allocators that do
> the same. There are also objects that may be used by multiple processes.

This is also true for normal user pages. And then, we do what memcg
does: first one to touch, gets accounted. I don't think deviating from
the memcg behavior for user pages makes much sense here.

A cache won't go away while it still have objects, even after the memcg
is removed (it is marked as dead)

> F.e what about shm?
>
>>>> /proc/slabinfo reflects this information, by listing the memcg-specific
>>>> slabs.
>>>
>>> What about /sys/kernel/slab/*?
>>
>> From the PoV of the global system, what you'll see is something like:
>> dentry , dentry(2:memcg1), dentry(2:memcg2), etc.
>
> Hmmm.. Would be better to have a hierachy there. /proc/slabinfo is more
> legacy.

I can take a look at that then. Assuming you agree with all the rest, is
looking into that a pre-requisite for merging, or is something that can
be deferred for a phase2 ? (We still don't do shrinkers, for instance,
so this is sure to have a phase2)