

(2012/04/21 6:57), 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.
>

Hmm, res_counter never goes naeative ?

> 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: Kamezawa Hiroyuki <kamezawa.hiroyu@jp.fujitsu.com>
> CC: Johannes Weiner <hannes@cmpxchg.org>

The code itself seems fine.

Reviewed-by: KAMEZAWA Hiroyuki <kamezawa.hiroyu@jp.fujitsu.com>
