Subject: Re: [PATCH v4 11/11] memcg: check memcg dirty limits in page writeback
Posted by Wu Fengguang on Sun, 31 Oct 2010 20:03:41 GMT
View Forum Message <> Reply to Message

On Sat, Oct 30, 2010 at 12:06:33AM +0800, Greg Thelen wrote:
> KAMEZAWA Hiroyuki <kamezawa.hiroyu@jp.fujitsu.com> writes:
>
> > On Fri, 29 Oct 2010 00:09:14 -0700
> > Greg Thelen <gthelen@google.com> wrote:
> >
> >> If the current process is in a non-root memcg, then
> >> balance_dirty_pages() will consider the memcg dirty limits
> >> as well as the system-wide limits. This allows different
> >> cgroups to have distinct dirty limits which trigger direct
> >> and background writeback at different levels.
> >>
> >> Signed-off-by: Andrea Righi <arighi@develer.com>
> >> Signed-off-by: Greg Thelen <gthelen@google.com>
> >
> > Acked-by: KAMEZAWA Hiroyuki <kamezawa.hiroyu@jp.fujitsu.com>

The "check both memcg&global dirty limit" looks much more sane than
the V3 implementation. Although it still has misbehaviors in some
cases, it's generally a good new feature to have.

Acked-by: Wu Fengguang <fengguang.wu@intel.com>

> > Ideally, I think some comments in the code for "why we need double-check system's
> > dirty limit and memcg's dirty limit" will be appreciated.
>
> I will add to the balance_dirty_pages() comment. It will read:
> /*
> * balance_dirty_pages() must be called by processes which are generating dirty
> * data. It looks at the number of dirty pages in the machine and will force
> * the caller to perform writeback if the system is over `vm_dirty_ratio'.
> ~~~~~~~~~~~~~~~~~~ ~~~~

To be exact, it tries to throttle the dirty speed so that
vm_dirty_ratio is not exceeded. In fact balance_dirty_pages() starts
throttling the dirtier slightly below vm_dirty_ratio.

> * If we're over `background_thresh' then the writeback threads are woken to
> * perform some writeout. The current task may have per-memcg dirty
> * limits, which are also checked.
> */

_____
Containers mailing list
Containers@lists.linux-foundation.org