Subject: Re: [PATCH] cgroup: Remove call to synchronize_rcu in cgroup_attach_task
Posted by Colin Cross on Sat, 22 Jan 2011 02:04:06 GMT

On Fri, Jan 21, 2011 at 5:17 PM, Bryan Huntsman <bryanh@codeaurora.org> wrote:
> On 11/23/2010 05:43 PM, Colin Cross wrote:
>> synchronize_rcu can be very expensive, averaging 100 ms in
>> some cases. In cgroup_attach_task, it is used to prevent
>> a task->cgroups pointer dereferenced in an RCU read side
>> critical section from being invalidated by delaying the call
>> to put_css_set until after an RCU grace period.
>>
>> To avoid the call to synchronize_rcu, make the put_css_set
>> call rcu-safe by moving the deletion of the css_set links
>> into rcu-protected free_css_set_rcu.
>>
>> The calls to check_for_release in free_css_set_rcu now occur
>> in softirq context, so convert all uses of the
>> release_list_lock spinlock to irq safe versions.
>>
>> The decrement of the cgroup refcount is no longer
>> synchronous with the call to put_css_set, which can result
>> in the cgroup refcount staying positive after the last call
>> to cgroup_attach_task returns. To allow the cgroup to be
>> deleted with cgroup_rmdir synchronously after
>> cgroup_attach_task, introduce a second refcount,
>> rmdir_count, that is decremented synchronously in
>> put_css_set. If cgroup_rmdir is called on a cgroup for
>> hich rmdir_count is zero but count is nonzero, reuse the
>> rmdir waitqueue to block the rmdir until the rcu callback
>> is called.
>>
>> Signed-off-by: Colin Cross <ccross@android.com>
>> ---
>>
>> This patch is similar to what you described. The main differences are
>> that I used a new atomic to handle the rmdir case, and I converted
>> check_for_release to be callable in softirq context rather than schedule
>> work in free_css_set_rcu. Your css_set scanning in rmdir sounds better,
>> I'll take another look at that. Is there any problem with disabling irqs
>> with spin_lock_irqsave in check_for_release?
>>
>> include/linux/cgroup.h | 6 ++
>> kernel/cgroup.c | 124 +++++++++++++++++++++++++++++++++--------------------
>> 2 files changed, 78 insertions(+), 52 deletions(-)
>>
>

> Colin, what became of this patch? I see this in your Tegra tree for
> Android.
>
>
http://android.git.kernel.org/?p=kernel/tegra.git;a=commit;h=05946a1e0fdb011ac0e6638ee60b18
1c584f127b
>
> I looked in linux-next but didn't see it there. This resolves a
> performance issue on MSM SMP so I'm curious if this is going upstream.
> Thanks.
>

It's been posted, there are no outstanding comments I am working on,
but they haven't been picked up.

_____