Subject: Re: [PATCH v2 06/11] memcg: kmem controller infrastructure
Posted by Greg Thelen on Mon, 13 Aug 2012 21:21:19 GMT
View Forum Message <> Reply to Message

On Mon, Aug 13 2012, Glauber Costa wrote:

>>
>> Here's the dmesg splat.
>>
>
> Do you always get this report in the same way?
> I managed to get a softirq inconsistency like yours, but the complaint
> goes for a different lock.

Yes, I repeatedly get the same dmesg splat below.

Once I your 'execute the whole memcg freeing in rcu callback' patch,
then the warnings are not printed. I'll take a closer look at the patch
soon.

>> [335.550398] ================================
>> [335.554739] [INFO: inconsistent lock state]
>> [335.559091] 3.5.0-dbg-DEV #3 Tainted: G   W
>> [335.563946] --------------------------------
>> [335.568290] inconsistent {SOFTIRQ-ON-W} -> {IN-SOFTIRQ-W} usage.
>> [335.574286] swapper/10/0 [HC0[0]:SC1[1]:HE1:SE0] takes:
>> [335.579508] (&(&rtpz->lock)->rlock){+.?...}, at: [<ffffffff8118216d>]
__mem_cgroup_free+0x8d/0x1b0
>> [335.588525] {SOFTIRQ-ON-W} state was registered at:
>> [335.593389]   [<ffffffff810cb073>] __lock_acquire+0x623/0x1a50
>> [335.599200]   [<ffffffff810cca55>] lock_acquire+0x95/0x150
>> [335.604670]   [<ffffffff81582531>] _raw_spin_lock+0x41/0x50
>> [335.610232]   [<ffffffff8118216d>] __mem_cgroup_free+0x8d/0x1b0
>> [335.616135]   [<ffffffff811822d5>] mem_cgroup_put+0x45/0x50
>> [335.621696]   [<ffffffff81182302>] mem_cgroup_destroy+0x22/0x30
>> [335.627592]   [<ffffffff810e093f>] cgroup_diput+0xbf/0x160
>> [335.633062]   [<ffffffff811a07ef>] d_delete+0x12f/0x1a0
>> [335.638276]   [<ffffffff8119671e>] vfs_rmdir+0x11e/0x140
>> [335.643565]   [<ffffffff81199173>] do_rmdir+0x113/0x130
>> [335.648773]   [<ffffffff8119a5e6>] sys_rmdir+0x16/0x20
>> [335.653900]   [<ffffffff8158c74f>] cstar_dispatch+0x7/0x1f
>> [335.659370] irq event stamp: 399732
>> [335.662846] hardirqs last enabled at (399732): [<ffffffff810e8e08>]
res_counter_uncharge_until+0x68/0xa0
>> [335.672383] hardirqs last disabled at (399731): [<ffffffff810e8dc8>]
res_counter_uncharge_until+0x28/0xa0
>> [335.681916] softirqs last enabled at (399710): [<ffffffff81085dd3>]
_local_bh_enable+0x13/0x20

```
>> [  335.690590] softirqs last disabled at (399711): [<ffffffff8158c48c>] call_softirq+0x1c/0x30
>> [  335.698914]
>> [  335.698914] other info that might help us debug this:
>> [  335.705415]  Possible unsafe locking scenario:
>> [  335.705415]
>> [  335.711317]        CPU0
>> [  335.713757]        ----
>> [  335.716198]   lock(&(&rtpz->lock)->rlock);
>> [  335.720282]   <Interrupt>
>> [  335.722896]     lock(&(&rtpz->lock)->rlock);
>> [  335.727153]
>> [  335.727153]  *** DEADLOCK ***
>> [  335.727153]
>> [  335.733055] no locks held by swapper/10/0.
>> [  335.737141]
>> [  335.737141] stack backtrace:
>> [  335.741483] Pid: 0, comm: swapper/10 Tainted: G        W    3.5.0-dbg-DEV #3
>> [  335.748510] Call Trace:
>> [  335.750952]  <IRQ>  [<ffffffff81579a27>] print_usage_bug+0x1fc/0x20d
>> [  335.757286]  [<ffffffff81058a9f>] ? save_stack_trace+0x2f/0x50
>> [  335.763098]  [<ffffffff810ca9ed>] mark_lock+0x29d/0x300
>> [  335.768309]  [<ffffffff810c9e10>] ? print_irq_inversion_bug.part.36+0x1f0/0x1f0
>> [  335.775599]  [<ffffffff810caffc>] __lock_acquire+0x5ac/0x1a50
>> [  335.781323]  [<ffffffff810cad34>] ? __lock_acquire+0x2e4/0x1a50
>> [  335.787224]  [<ffffffff8118216d>] ? __mem_cgroup_free+0x8d/0x1b0
>> [  335.793212]  [<ffffffff810cca55>] lock_acquire+0x95/0x150
>> [  335.798594]  [<ffffffff8118216d>] ? __mem_cgroup_free+0x8d/0x1b0
>> [  335.804581]  [<ffffffff810e8ddd>] ? res_counter_uncharge_until+0x3d/0xa0
>> [  335.811263]  [<ffffffff81582531>] _raw_spin_lock+0x41/0x50
>> [  335.816731]  [<ffffffff8118216d>] ? __mem_cgroup_free+0x8d/0x1b0
>> [  335.822724]  [<ffffffff8118216d>] __mem_cgroup_free+0x8d/0x1b0
>> [  335.828538]  [<ffffffff811822d5>] mem_cgroup_put+0x45/0x50
>> [  335.834002]  [<ffffffff811828a6>] __memcg_kmem_free_page+0xa6/0x110
>> [  335.840256]  [<ffffffff81138109>] free_accounted_pages+0x99/0xa0
>> [  335.846243]  [<ffffffff8107b09f>] free_task+0x3f/0x70
>> [  335.851278]  [<ffffffff8107b18c>] __put_task_struct+0xbc/0x130
>> [  335.857094]  [<ffffffff81081524>] delayed_put_task_struct+0x54/0xd0
>> [  335.863338]  [<ffffffff810fd354>] __rcu_process_callbacks+0x1e4/0x490
>> [  335.869757]  [<ffffffff810fd62f>] rcu_process_callbacks+0x2f/0x80
>> [  335.875835]  [<ffffffff810862f5>] __do_softirq+0xc5/0x270
>> [  335.881218]  [<ffffffff810c49b4>] ? clockevents_program_event+0x74/0x100
>> [  335.887895]  [<ffffffff810c5d94>] ? tick_program_event+0x24/0x30
>> [  335.893882]  [<ffffffff8158c48c>] call_softirq+0x1c/0x30
>> [  335.899179]  [<ffffffff8104cefd>] do_softirq+0x8d/0xc0
>> [  335.904301]  [<ffffffff810867de>] irq_exit+0xae/0xe0
>> [  335.909251]  [<ffffffff8158cc3e>] smp_apic_timer_interrupt+0x6e/0x99
>> [  335.915591]  [<ffffffff8158ba9c>] apic_timer_interrupt+0x6c/0x80
>> [  335.921583]  <EOI>  [<ffffffff810530e7>] ? default_idle+0x67/0x270
```

>> [335.927741] [<ffffffff810530e5>] ? default_idle+0x65/0x270
>>