Subject: Re: [PATCH 2/2] Make res_counter hierarchical
Posted by yamamoto on Wed, 12 Mar 2008 23:36:50 GMT
View Forum Message <> Reply to Message

> > @@ -36,10 +37,26 @@ int res_counter_charge(struct res_counter *counter, unsigned long val)
> > {
> > int ret;
> > unsigned long flags;
> > + struct res_counter *c, *unroll_c;
> > +
> > + local_irq_save(flags);
> > + for (c = counter; c != NULL; c = c->parent) {
> > + spin_lock(&c->lock);
> > + ret = res_counter_charge_locked(c, val);
> > + spin_unlock(&c->lock);
> > + if (ret < 0)
> > + goto unroll;
> > + }
> > + local_irq_restore(flags);
> > + return 0;
> >
> > - spin_lock_irqsave(&counter->lock, flags);
> > - ret = res_counter_charge_locked(counter, val);
> > - spin_unlock_irqrestore(&counter->lock, flags);
> > +unroll:
> > + for (unroll_c = counter; unroll_c != c; unroll_c = unroll_c->parent) {
> > + spin_lock(&unroll_c->lock);
> > + res_counter_uncharge_locked(unroll_c, val);
> > + spin_unlock(&unroll_c->lock);
> > + }
> > + local_irq_restore(flags);
> > return ret;
> > }
>
> what prevents the topology (in particular, ->parent pointers) from
> changing behind us?
>
> YAMAMOTO Takashi

to answer myself: cgroupfs rename doesn't allow topological changes
in the first place.

btw, i think you need to do the same for res_counter_limit_check_locked
as well. i'm skeptical about doing these complicated stuffs in kernel,
esp. in this potentially performance critical code.

YAMAMOTO Takashi

Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers