

---

Subject: Re: [RFC] [PATCH] memory controller background reclamation  
Posted by [Balbir Singh](#) on Mon, 26 Nov 2007 03:00:19 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

YAMAMOTO Takashi wrote:

```
> hi,
>
>>> --- linux-2.6.24-rc2-mm1-kame-pd/kernel/res_counter.c.BACKUP 2007-11-14
16:05:52.000000000 +0900
>>> +++ linux-2.6.24-rc2-mm1-kame-pd/kernel/res_counter.c 2007-11-22 15:14:32.000000000
+0900
>>> @@ -17,6 +17,8 @@ void res_counter_init(struct res_counter
>>> {
>>> spin_lock_init(&counter->lock);
>>> counter->limit = (unsigned long long)LLONG_MAX;
>>> + counter->high_watermark = (unsigned long long)LLONG_MAX;
>>> + counter->low_watermark = (unsigned long long)LLONG_MAX;
>> Should low watermark also be LLONG_MAX?
>
> what else do you suggest? 0?
```

Something invalid or a good default value. I think LLONG\_MAX  
is good for now, that ensures that no background reclaim  
happens till the administrator sets it up.

```
> currently it doesn't matter much because low_watermark is not used at all
> as far as high_watermark is LLONG_MAX.
>
```

Don't we use by checking `res_counter_below_low_watermark()`?

```
>>> +static void
>>> +mem_cgroup_reclaim(struct work_struct *work)
>>> +{
>>> + struct mem_cgroup * const mem =
>>> + container_of(work, struct mem_cgroup, reclaim_work);
>>> + int batch_count = 128; /* XXX arbitrary */
>> Could we define and use something like MEM_CGROUP_BATCH_COUNT for now?
>> Later we could consider and see if it needs to be tunable. numbers are
>> hard to read in code.
>
> although i don't think it makes sense, i can do so if you prefer.
>
```

Using numbers like 128 make the code unreadable. I prefer something  
like MEM\_CGROUP\_BATCH\_COUNT since its more readable than 128. If we ever  
propagate batch\_count to other dependent functions, I'd much rather do  
it with a well defined name.

```
>>> +
>>> + for (; batch_count > 0; batch_count--) {
>>> + if (res_counter_below_low_watermark(&mem->res))
>>> + break;
>> Shouldn't we also check to see that we start reclaim in background only
>> when we are above the high watermark?
>
> i don't understand what you mean. can you explain?
> highwatermark is checked by mem_cgroup_charge_common before waking
> these threads.
>
```

OK, that clarifies

```
>> I'll start some tests on these patches.
>
> thanks.
>
> YAMAMOTO Takashi
```

--

Warm Regards,  
Balbir Singh  
Linux Technology Center  
IBM, ISTL

---

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

---