On Mon, 27 Feb 2012 14:58:47 -0800
Suleiman Souhlal <ssouhlal@FreeBSD.org> wrote:

> This is used to indicate that we don't want an allocation to be accounted
> to the current cgroup.
>
> Signed-off-by: Suleiman Souhlal <suleiman@google.com>

I don't like this.

Please add

___GFP_ACCOUNT "account this allocation to memcg"

Or make this as slab's flag if this work is for slab allocation.

Thanks,
-Kame

> ---
> include/linux/gfp.h | 2 ++
> 1 files changed, 2 insertions(+), 0 deletions(-)
>
> diff --git a/include/linux/gfp.h b/include/linux/gfp.h
> index 581e74b..765c20f 100644
> --- a/include/linux/gfp.h
> +++ b/include/linux/gfp.h
> @@ -23,6 +23,7 @@ struct vm_area_struct;
> #define ___GFP_REPEAT 0x400u
> #define ___GFP_NOFAIL 0x800u
> #define ___GFP_NORETRY 0x1000u
> +#define ___GFP_NOACCOUNT 0x2000u
> #define ___GFP_COMP 0x4000u
> #define ___GFP_ZERO 0x8000u
> #define ___GFP_NOMEMALLOC 0x10000u
> @@ -76,6 +77,7 @@ struct vm_area_struct;
> #define __GFP_REPEAT ((__force gfp_t)___GFP_REPEAT) /* See above */
> #define __GFP_NOFAIL ((__force gfp_t)___GFP_NOFAIL) /* See above */
> #define __GFP_NORETRY ((__force gfp_t)___GFP_NORETRY) /* See above */
> +#define __GFP_NOACCOUNT ((__force gfp_t)___GFP_NOACCOUNT) /* Don't account to the
current cgroup */
> #define __GFP_COMP ((__force gfp_t)___GFP_COMP) /* Add compound page metadata */

> #define __GFP_ZERO ((__force gfp_t)___GFP_ZERO) /* Return zeroed page on success */
> #define __GFP_NOMEMALLOC ((__force gfp_t)___GFP_NOMEMALLOC) /* Don't use emergency reserves */
> --
> 1.7.7.3
>
> --
> To unsubscribe, send a message with 'unsubscribe linux-mm' in
> the body to majordomo@kvack.org. For more info on Linux MM,
> see: http://www.linux-mm.org/ .
> Fight unfair telecom internet charges in Canada: sign http://stopthemeter.ca/
> Don't email: email@kvack.org
>

Subject: Re: [PATCH 04/10] memcg: Introduce __GFP_NOACCOUNT.
Posted by Glauber Costa on Wed, 29 Feb 2012 16:53:28 GMT
View Forum Message <> Reply to Message

On 02/29/2012 03:00 AM, KAMEZAWA Hiroyuki wrote:
> On Mon, 27 Feb 2012 14:58:47 -0800
> Suleiman Souhlal<ssouhlal@FreeBSD.org> wrote:
>
>> This is used to indicate that we don't want an allocation to be accounted
>> to the current cgroup.
>>
>> Signed-off-by: Suleiman Souhlal<suleiman@google.com>
>
> I don't like this.
>
> Please add
>
> ___GFP_ACCOUNT "account this allocation to memcg"
>
> Or make this as slab's flag if this work is for slab allocation.
>
> Thanks,
> -Kame
>
>
+1.

This is enough to replace the registration I originally proposed, and I
do think it'll do us good, despite the churn drawback of having to go
patching stuff everywhere.

Subject: Re: [PATCH 04/10] memcg: Introduce __GFP_NOACCOUNT.
Posted by Suleiman Souhlal on Wed, 29 Feb 2012 19:09:50 GMT

View Forum Message <> Reply to Message

On Tue, Feb 28, 2012 at 10:00 PM, KAMEZAWA Hiroyuki
<kamezawa.hiroyu@jp.fujitsu.com> wrote:
> On Mon, 27 Feb 2012 14:58:47 -0800
> Suleiman Souhlal <ssouhlal@FreeBSD.org> wrote:
>
>> This is used to indicate that we don't want an allocation to be accounted
>> to the current cgroup.
>>
>> Signed-off-by: Suleiman Souhlal <suleiman@google.com>
>
> I don't like this.
>
> Please add
>
> ___GFP_ACCOUNT "account this allocation to memcg"
>
> Or make this as slab's flag if this work is for slab allocation.

We would like to account for all the slab allocations that happen in
process context.

Manually marking every single allocation or kmem_cache with a GFP flag
really doesn't seem like the right thing to do..

Can you explain why you don't like this flag?

-- Suleiman