
Subject: Re: [RFC][PATCH 3/7] Data structures changes for RSS accounting
Posted by [xemul](#) on Tue, 13 Mar 2007 07:10:43 GMT

[View Forum Message](#) <> [Reply to Message](#)

Dave Hansen wrote:

> On Mon, 2007-03-12 at 20:19 +0300, Pavel Emelianov wrote:
>> Dave Hansen wrote:
>>> On Mon, 2007-03-12 at 19:16 +0300, Kirill Korotaev wrote:
>>>> now VE2 maps the same page. You can't determine whether this page is mapped
>>>> to this container or another one w/o page->container pointer.
>>> Hi Kirill,
>>>
>>> I thought we can always get from the page to the VMA. rmap provides
>>> this to us via page->mapping and the 'struct address_space' or anon_vma.
>>> Do we agree on that?
>> Not completely. When page is unmapped from the *very last*
>> user its *first* toucher may already be dead. So we'll never
>> find out who it was.
>
> OK, but this is assuming that we didn't *un*account for the page when
> the last user of the "owning" container stopped using the page.

That's exactly what we agreed on during our discussions:
When page is get touched it is charged to this container.
When page is get touched again by new container it is NOT
charged to new container, but keeps holding the old one
till it (the page) is completely freed. Nobody worried the
fact that a single page can hold container for good.

OpenVZ beancounters work the other way (and we proposed this
solution when we first sent the patches). We keep track of
all the containers (i.e. beancounters) holding this page.

>>> We can also get from the vma to the mm very easily, via vma->vm_mm,
>>> right?
>>>
>>> We can also get from a task to the container quite easily.
>>>
>>> So, the only question becomes whether there is a 1:1 relationship
>>> between mm_structs and containers. Does each mm_struct belong to one
>> No. The question is "how to get a container that touched the
>> page first" which is the same as "how to find mm_struct which
>> touched the page first". Obviously there's no answer on this
>> question unless we hold some direct page->container reference.
>> This may be a hash, a direct on-page pointer, or mirrored
>> array of pointers.
>
> Or, you keep track of when the last user from the container goes away,

> and you effectively account it to another one.

We can migrate page to another user but we decided to implement it later after accepting simple accounting.

> Are there problems with shifting ownership around like this?

>

> -- Dave

>

>

Containers mailing list

Containers@lists.osdl.org

<https://lists.osdl.org/mailman/listinfo/containers>
