
Subject: Re: [RFC][PATCH 2/7] RSS controller core
Posted by [Dave Hansen](#) on Tue, 13 Mar 2007 17:26:33 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Mon, 2007-03-12 at 22:04 -0800, Andrew Morton wrote:

> So these mmaped pages will continue to be shared across all guests. The
> problem boils down to "which guest(s) get charged for each shared page".
>
> A simple and obvious and easy-to-implement answer is "the guest which paged
> it in". I think we should firstly explain why that is insufficient.

My first worry was that this approach is unfair to the poor bastard that happened to get started up first. If we have a bunch of containerized web servers, the poor guy who starts Apache first will pay the price for keeping it in memory for everybody else.

That said, I think this is naturally worked around. The guy charged unfairly will get reclaim started on himself sooner. This will tend to page out those pages that he was being unfairly charged for. Hopefully, they will eventually get pretty randomly (eventually evenly) spread among all users. We just might want to make sure that we don't allow ptes (or other new references) to be re-established to pages like this when we're trying to reclaim them. Either that, or force the next toucher to take ownership of the thing. But, that kind of arbitrary ownership transfer can't happen if we have rigidly defined boundaries for the containers.

The other concern is that the memory load on the system doesn't come from the first user ("the guy who paged it in"). The long-term load comes from "the guy who keeps using it." The best way to exemplify this is somebody who read(s) a page in, followed by another guy mmap()ing the same page. The guy who did the read will get charged, and the mmap()er will get a free ride. We could probably get an idea when this kind of stuff is happening by comparing page->count and page->_mapcount, but it certainly wouldn't be conclusive. But, does this kind of nonsense even happen in practice?

-- Dave

Containers mailing list
Containers@lists.osdl.org
<https://lists.osdl.org/mailman/listinfo/containers>
