
Subject: Re: Linux-VServer example results for sharing vs. separate mappings ...

Posted by [Ethan Solomita](#) on Wed, 28 Mar 2007 01:22:02 GMT

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

Herbert Poetzl wrote:

> On Sat, Mar 24, 2007 at 12:19:06PM -0800, Andrew Morton wrote:

>

>> Or change the reclaim code so that a page which hasn't

>> been referenced from a process within its hardware

>> container is considered unreferenced (so it gets reclaimed).

>>

>

> that might easily lead to some ping-pong behaviour,

> when two similar guest are executing similar binaries

> but not at the same time ...

>

It might lead to that, but I don't think it would become pathological "easily". If a system has been up for a long time, it's easy to image pagecache pages lying everywhere just because someone somewhere is still using them.

I suggest a variant on what Andrew says: don't change reclaim. Instead, when referencing a page, don't mark the page as referenced if the current task is not permitted to allocate from the page's node. I'm thinking in terms of cpusets, with each task having a nodemask of `mems_allowed`. This may result in a page being thrown out unnecessarily and brought back in from disk, but when memory is tight that is what happens. An optimization might be to keep track of who is referencing the page and migrate it to their memory instead of reclaiming it, but that would require reclaim to know the task/cpuset/container of the referencing task.

-- Ethan

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>
