Subject: Re: controlling mmap()'d vs read/write() pages
Posted by Nick Piggin on Fri, 23 Mar 2007 10:47:45 GMT

View Forum Message <> Reply to Message

Eric W. Biederman wrote:
> Nick Piggin <nickpiggin@yahoo.com.au> writes:
>
>
>>Eric W. Biederman wrote:
>>
>>>Dave Hansen <hansendc@us.ibm.com> writes:
>>>
>>>
>>>
>>>>So, I think we have a difference of opinion. I think it's _all_ about
>>>>memory pressure, and you think it is _not_ about accounting for memory
>>>>pressure. :) Perhaps we mean different things, but we appear to
>>>>disagree greatly on the surface.
>>>
>>>
>>>I think it is about preventing a badly behaved container from having a
>>>significant effect on the rest of the system, and in particular other
>>>containers on the system.
>>
>>That's Dave's point, I believe. Limiting mapped memory may be
>>mostly OK for well behaved applications, but it doesn't do anything
>>to stop bad ones from effectively DoSing the system or ruining any
>>guarantees you might proclaim (not that hard guarantees are always
>>possible without using virtualisation anyway).
>>
>>This is why I'm surprised at efforts that go to such great lengths
>>to get accounting "just right" (but only for mmaped memory). You
>>may as well not even bother, IMO.
>>
>>Give me an RSS limit big enough to run a couple of system calls and
>>a loop...
>
>
> Would any of them work on a system on which every filesystem was on
> ramfs, and there was no swap? If not then they are not memory attacks
> but I/O attacks.
>
> I completely concede that you can DOS the system with I/O if that is
> not limited as well.
>
> My point is that is not a memory problem but a disk I/O problem which is
> much easier to and cheaper to solve. Disk I/O is fundamentally a slow
> path which makes it hard to modify it in a way that negatively affects

> system performance.
>
> I don't think with a memory RSS limit you can DOS the system in a way
> that is purely about memory. You have to pick a different kind of DOS
> attack.

It can be done trivially without performing any IO or swap, yes.

--
SUSE Labs, Novell Inc.
Send instant messages to your online friends http://au.messenger.yahoo.com

_____