
Subject: Re: [RFC][PATCH 2/7] RSS controller core
Posted by [Herbert Poetzl](#) on Tue, 13 Mar 2007 15:11:55 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Tue, Mar 13, 2007 at 06:10:55PM +0300, Kirill Korotaev wrote:

> >> So what to do when virtual physical limit is hit?
> >> OOM-kill current task?
> >
> >
> > when the RSS limit is hit, but there _are_ enough
> > pages left on the physical system, there is no
> > good reason to swap out the page at all
> >
> > - there is no benefit in doing so (performance
> > wise, that is)
> >
> > - it actually hurts performance, and could
> > become a separate source for DoS
> >
> > what should happen instead (in an ideal world :)
> > is that the page is considered swapped out for
> > the guest (add guest penalty for swapout), and
> > when the page would be swapped in again, the guest
> > takes a penalty (for the 'virtual' page in) and
> > the page is returned to the guest, possibly kicking
> > out (again virtually) a different page
>
> great. I agree with that.

> Just curious why current vserver code kills arbitrary
> task in container then?

because it obviously lacks the finesse of OpenVZ code :)

seriously, handling the OOM kills inside a container
has never been a real world issue, as once you are
really out of memory (and OOM starts killing) you
usually have lost the game anyways (i.e. a guest restart
or similar is required to get your services up and
running again) and OOM killer decisions are not perfect
in mainline either, but, you've probably seen the
FIXME and TODO entries in the code showing that this
is work in progress ...

> >>> - accounting and limits have to be consistent
> >>> and should roughly represent the actual used
> >>> memory/swap (modulo optimizations, I can go
> >>> into detail here, if necessary)

> >>
> >>This is true for current implementation for
> >>booth - this patchset and OpenVZ beancounters.
> >>
> >>If you sum up the physpages values for all containers
> >>you'll get the exact number of RAM pages used.
> >
> >
> > hmm, including or excluding the host pages?
>
> depends on whether you will include beanocunter 0 usages or not :)

so that is an option then?

best,
Herbert

> Kirill

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