
Subject: Re: Re: [ckrm-tech] [PATCH] BC: resource beancounters (v4) (added user memory)

Posted by [Rohit Seth](#) on Tue, 12 Sep 2006 00:28:58 GMT

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

On Mon, 2006-09-11 at 23:48 +0400, Kir Kolyshkin wrote:

> Rohit Seth wrote:

> > On Mon, 2006-09-11 at 11:25 -0700, Chandra Seetharaman wrote:

> >

> >> On Fri, 2006-09-08 at 14:43 -0700, Rohit Seth wrote:

> >> <snip>

> >>

> >>

> >>>> Guarantee may be one of

> >>>>

> >>>> 1. container will be able to touch that number of pages

> >>>> 2. container will be able to sys_mmap() that number of pages

> >>>> 3. container will not be killed unless it touches that number of pages

> >>>> 4. anything else

> >>>>

> >>>> I would say (1) with slight modification

> >>>> "container will be able to touch _at least_ that number of pages"

> >>>>

> >>>>

> >>> Does this scheme support running of tasks outside of containers on the

> >>> same platform where you have tasks running inside containers. If so

> >>> then how will you ensure processes running outside any container will

> >>> not leave less than the total guaranteed memory to different containers.

> >>>

> >>>

> >> There could be a default container which doesn't have any guarantee or
> >> limit.

> >>

> >

> > First, I think it is critical that we allow processes to run outside of

> > any container (unless we know for sure that the penalty of running a

> > process inside a container is very very minimal).

> >

> (1) there is a set of processes running outside of any container. In

> OpenVZ we call that "VE0" or "host system", probably Chandra meant that

> by "default container".

> (2) The host system is used to manage the containers (start/stop/set

> parameters/create/destroy).

> (3) the penalty of running a process inside a container is indeed very low.

>

> > And anything running outside a container should be limited by default

> > Linux settings.

> >

> (4) due to (2), it is not recommended to run anything but the tasks used
> to manage the containers -- otherwise your gonna have security problems

Just like you want to run those special threads outside of any container, some sysadmin might be interested in running different processes that they don't want to bind to any container limits.

I think it is critical that you provide the capability to have tasks running outside any container. Whether sysadmin wants to do it or not for a system is a different thing.

-rohit
