Subject: Re: [patch00/05]: Containers(V2)- Introduction
Posted by Rohit Seth on Wed, 20 Sep 2006 18:57:40 GMT
View Forum Message <> Reply to Message

On Wed, 2006-09-20 at 20:37 +0200, Peter Zijlstra wrote:
> On Wed, 2006-09-20 at 10:50 -0700, Rohit Seth wrote:
> > On Thu, 2006-09-21 at 03:00 +1000, Nick Piggin wrote:
> > > (this time to the lists as well)
> > >
> > > Peter Zijlstra wrote:
> > >
> > > > I'd much rather containterize the whole reclaim code, which should not
> > > > be too hard since he already adds a container pointer to struct page.
> > >
> > >
> > >
> >
> > Right now the memory handler in this container subsystem is written in
> > such a way that when existing kernel reclaimer kicks in, it will first
> > operate on those (container with pages over the limit) pages first. But
> > in general I like the notion of containerizing the whole reclaim code.
>
> Patch 5/5 seems to have a horrid deactivation scheme.
>
> > > > I still have to reread what Rohit does for file backed pages, that gave
> > > > my head a spin.
> >
> > Please let me know if there is any specific part that isn't making much
> > sense.
>
> Well, the whole over the limit handler is quite painfull, having taken a
> second reading it isn't all that complex after all, just odd.
>

It is very basic right now.

> You just start invalidating whole files for file backed pages. Granted,
> this will get you below the threshold. but you might just have destroyed
> your working set.
>

When a container gone over the limit then it is okay to penalize it. I agree that I'm not making an attempt to maintain the current working set. Any suggestions that I can incorporate to improve this algorithm will be very appreciated.

> Pretty much the same for you anonymous memory handler, you scan through
> the pages in linear fashion and demote the first that you encounter.
>

> Both things pretty thoroughly destroy the existing kernel reclaim.
>

I agree that with in a container I need to do add more smarts to (for example) not do a linear search. Simple additions like last task or last mapping visited could be useful. And I definitely want to improve on that.

Though it should not destroy the existing kernel reclaim. Pages belonging to over the limit container should be the first ones to either get flushed out to FS or swapped if necessary. (Means that is the cost that you will have to pay if you, for example, want to container your tar to 100MB memory foot print).

-rohit