Subject: Re: [ckrm-tech] [RFC][PATCH 5/7] UBC: kernel memory accounting (core)
Posted by dev on Mon, 21 Aug 2006 10:38:17 GMT
View Forum Message <> Reply to Message

Dave Hansen wrote:
> On Fri, 2006-08-18 at 13:31 +0400, Kirill Korotaev wrote:
>
>>they all are troublesome :/
>>user can create lots of vmas, w/o page tables.
>>lots of fdsets, ipcids.
>>These are not reclaimable.
>
>
> I guess one of my big questions surrounding these patches is why the
> accounting is done with pages.
probably you missed patch details a bit.
accounting is done:
1. in pages for objects allocated by buddy allocator
2. in slabs for objects allocated from caches

> If there really is a need to limit these
> different kernel objects, then why not simply write patches to limit
> *these* *objects*? I trust there is a very good technical reason for
> doing this, I just don't understand why, yet.
The one reason is that such an accounting allows to estimate the memory
used/required by containers, while limitations by objects:
- per object accounting/limitations do not provide any memory estimation
- having a big number of reasonably high limits still allows the user
  to consume big amount of memory. I.e. the sum of all the limits tend
  to be high and potentially DoS exploitable :/
- memory is easier to setup/control from user POV.
  having hundreds of controls is good, but not much user friendly.

Thanks,
Kirill