Subject: Re: [RFC][PATCH 5/7] UBC: kernel memory accounting (core)
Posted by dev on Fri, 18 Aug 2006 09:36:19 GMT
View Forum Message <> Reply to Message

Rohit Seth wrote:
> On Thu, 2006-08-17 at 17:27 +0400, Kirill Korotaev wrote:
>
>>>If I'm reading this patch right then seems like you are making page
>>>allocations to fail w/o (for example) trying to purge some pages from
>>>the page cache belonging to this container. Or is that reclaim going to
>>>come later?
>>
>>charged kernel objects can't be _reclaimed_. how do you propose
>>to reclaim tasks page tables or files or task struct or vma or etc.?
>
>
>
> I agree that kernel objects cann't be reclaimed easily. But what you
> are proposing is also not right. Returning failure w/o doing any
> reclaim on pages (that are reclaimable) is not useful. And this is why
> I asked, is this change going to be part of next set of patches (as
> current set of patches are only tracking kernel usage).
1. reclaiming user resources is not that good idea as it looks to you.
such solutions end up with lots of resources spent on reclaim.
for user memory reclaims mean consumption of expensive disk I/O bandwidth
which reduces overall system throughput and influences other users.

2. kernel memory is mostly not reclaimable. can you reclaim vma structs or ipc ids?
even with page tables it is not that easy.
And the fact is:
 - kernel memory consumtion is usually less then user memory,
   so it's not worth reclaiming it.
 - reclaiming can result in kind of user service deadlocks when you are
   unable to handle user requests gracefully anymore.
   See my email to Dave with an example.
 - our solution _is_ right and works (for >3 years in production already).
   If desired it _can_ be extended with reclamation.

Kirill