

---

Subject: Re: Linux-VServer example results for sharing vs. separate mappings ...  
Posted by [Balbir Singh](#) on Mon, 26 Mar 2007 02:36:07 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Andrew Morton wrote:

>> Don't we break the global LRU with this scheme?  
>  
> Sure, but that's deliberate!  
>  
> (And we don't have a global LRU - the LRUs are per-zone).  
>

Yes, true. But if we use zones for containers and say we have 400 of them, with all of them under limit. When the system wants to reclaim memory, we might not end up reclaiming the best pages. Am I missing something?

>>> b) Create a new memory abstraction, call it the "software zone", which  
>>> is mostly decoupled from the present "hardware zones". Most of the MM  
>>> is reworked to use "software zones". The "software zones" are  
>>> runtime-resizeable, and obtain their pages via some means from the  
>>> hardware zones. A container uses a software zone.  
>>>  
>> I think the problem would be figuring out where to allocate memory from?  
>> What happens if a software zone spans across many hardware zones?  
>  
> Yes, that would be the tricky part. But we generally don't care what  
> physical zone user pages come from, apart from NUMA optimisation.  
>  
>> The reclaim mechanism proposed \*does not impact the non-container users\*.  
>  
> Yup. Let's keep plugging away with Pavel's approach, see where it gets us.  
>

Yes, we have some changes that we've made to the reclaim logic, we hope to integrate a page cache controller soon. We are also testing the patches. Hopefully soon enough, they'll be in a good state and we can request you to merge the containers and the rss limit (plus page cache) controller soon.

--

Warm Regards,  
Balbir Singh  
Linux Technology Center  
IBM, ISTL

---

Containers mailing list

