

(this time to the lists as well)

Peter Zijlstra wrote:

- > I'd much rather containerize the whole reclaim code, which should not
- > be too hard since he already adds a container pointer to struct page.

Yes, and I tend to agree with you. I probably wasn't clear, but I was mainly talking about just the memory resource tracking part of this patchset.

I am less willing to make a judgement about reclaim, because I don't know very much about the workloads or the guarantees they attempt to provide.

- > Esp. when we get some of my page reclaim abstractions merged, moving the
- > reclaim from struct zone to a container is not a lot of work. (this is
- > basically what one of the ckrn mm policies did too)

I do agree that it would be nicer to not have a completely different scheme for doing their own page reclaim, but rather use the existing code (*provided* that it is designed in the same, minimally intrusive manner as the page tracking).

I can understand how it is attractive to create a new subsystem to solve a particular problem, but once it is in the kernel it has to be maintained regardless, so if it can be done in a way that shares more of the current infrastructure (nicely) then that would be a better solution.

I like that they're investigating the use of memory nodes for this. It seems like the logical starting place.

- > I still have to reread what Rohit does for file backed pages, that gave
- > my head a spin.
- > I've been thinking a bit on that problem, and it would be possible to
- > share all address_space pages equally between attached containers, this
- > would lose some accuracy, since one container could read 10% of the file
- > and another 90%, but I don't think that is a common scenario.

Yeah, I'm not sure about that. I don't think really complex schemes

are needed... but again I might need more knowledge of their workloads and problems.

--

SUSE Labs, Novell Inc.

Send instant messages to your online friends <http://au.messenger.yahoo.com>
