Subject: Re: How Inactive may be much greather than cached?
Posted by Nick Piggin on Thu, 18 Oct 2007 07:27:00 GMT
View Forum Message <> Reply to Message

On Thursday 18 October 2007 17:14, Vasily Averin wrote:
> Nick Piggin wrote:
> > Hi,
> >
> > On Thursday 18 October 2007 16:24, Vasily Averin wrote:
> >> Hi all,
> >>
> >> could anybody explain how "inactive" may be much greater than "cached"?
> >> stress test (http://weather.ou.edu/~apw/projects/stress/) that writes
> >> into removed files in cycle puts the node to the following state:
> >>
> >> MemTotal: 16401648 kB
> >> MemFree: 636644 kB
> >> Buffers: 1122556 kB
> >> Cached: 362880 kB
> >> SwapCached: 700 kB
> >> Active: 1604180 kB
> >> Inactive: 13609828 kB
> >>
> >> At the first glance memory should be freed on file closing, nobody
> >> refers to file and ext3_delete_inode() truncates inode. We can see that
> >> memory is go away from "cached", however could somebody explain why it
> >> become "invalid" instead be freed? Who holds the references to these
> >> pages?
> >
> > Buffers, swap cache, and anonymous.
>
> But buffers and swap cache are low (1.1 Gb and 700kB in this example) and
> anonymous should go away when process finished.

Ah, I didn't see it was an order of magnitude out.

Some filesystems, including I believe, ext3 with data=ordered,
can leave orphaned pages around after they have been truncated
out of the pagecache. These pages get left on the LRU and vmscan
reclaims them pretty easily.

Try ext3 data=writeback, or even ext2.