

Re: No swap partition ?

Source: <http://linux.derkeiler.com/Newsgroups/comp.os.linux.development.system/2005-01/0251.html>

From: Kasper Dupont (kasperd_at_daimi.au.dk)

Date: 01/18/05

Date: Tue, 18 Jan 2005 08:19:48 +0100

Robert Redelmeier wrote:

>

> *Kasper Dupont <kasperd@daimi.au.dk> wrote:*

>

>> *But why remove cached pages you are going to need again*

>> *soon? There will be process data that is never going to be*

>> *used, why do you want to keep that data in RAM?*

>

> *Many cached pages are never needed again, or need to be*

> *written comparatively early.*

Of course you won't need all the cached pages. Still the more RAM you can use for caching, the more likely a particular page is to be in cache, once you need it.

>

>> *In most cases that would slow down the system, but since*

>> *the NFS server is implemented in the kernel, that is not*

>> *going to be too much of a problem.*

>

> *Codepage reloads are no slower than swap, and considerably*

> *less complex.*

Did you ever look on the code? Finding a page on swap is trivial. Finding a page of a memory mapped file means you need to go down through the file system specific drivers and call the get block for each block (or is it sector?) in the page, which again means you will need to read file system meta data. The complexity of this depends on the particular file system in use.

>

>>> *Or if big, the nfsd will have to recognize memory pressure & react.*

>> *React how?*

>

> *A monolithic NFSD would hold memory and have to start writing*

> *out and/or clearing cache.*

Re: No swap partition ?

That is definitely not the job of nfs.

- > *Some people think they can do a*
- > *better job of running cache than the kernel.*

You seem to be one of those people.

- >
- > > *Do I have a choice? You are the one who will not let*
- > > *the kernel do its job, but will rather force it to keep*
- > > *particular data in RAM all the time.*
- >
- > *That's one way of looking at it. I prefer to make the*
- > *poor overtaxed kernel's job easier.*

Unless you actually remove all of that code from the kernel, that argument doesn't hold.

- >
- > > *Can you really strip it down to just 2MB?*
- >
- > *Somewhere around there. I remember my first Linux Slackware 3.3*
- > *on i486 with 4 MB didn't need swap (but it helped).*

But is that really what you use on your NFS servers?

- >
- > > *On my system even init have more than 1MB of anonymous*
- >
- > *Sure. That doesn't mean it gets used.*

What do you mean by gets used? Only if the pages were allocated and never written they will be CoW of the zero page. If they have been written and then never gets used again, I don't want them to be in RAM. And there surely are pages that you can write to swap and forget.

- >
- > > *mappings. You can easilly find out, because you can just*
- > > *create a 3MB swap file/partition and see if it gets filled.*
- >
- > *Or you can boot with a restrictive bootparm like mem=6M*

That wouldn't answer the question. The kernel needs memory for other stuff as well. And you need some amount of memory for caching memory mappings in order for the start scripts to actually execute.

- >
- > > *IMHO the minimum for an NFS server would be: init, syslogd,*

Re: No swap partition ?

comp.os.linux.development.system: Re: No swap partition ?

> > *klogd, ntpd, sshd, gpm, mingetty*
>
> *Sounds like a reasonable list. I don't know of any util that*
> *will show the minimum CoW usage of these.*

The VmRSS field would be of some interest. Maybe VmData would also be of some interest, though I'm not really sure it means what the name suggests.

--
Kasper Dupont