

[Index](#)

[Rules](#)

[Search](#)

[Register](#)

[Login](#)

#1 2014-07-23 01:33:55

davidm

Member

Registered: 2009-04-25

Posts: 366

I thought I would get some input on these settings from knowledgeable Arch Linux users.

vm.swappiness (tendency to use swap, 0 = prevent OOM only, 100 = swap often)

vm.vfs_cache_pressure (tendency to reclaim swap space back to memory? default 100?)

vm.min_free_kbytes (physical memory to try to keep as reserve to prevent OOM, default depends on RAM)

I have mine set to:

```
vm.swappiness=50
vm.vfs_cache_pressure=400
vm.min_free_kbytes=262144
```

I'm using btrfs and have a loopback swap file.

6 GB Physical RAM
Swap ~3GB

```
# free -m
```

	total	used	free	shared	buffers	cached	
Mem:		5906	4376	1530	66	0	2830
-/+ buffers/cache:			1545	4361			
Swap:		3071	142	2929			

This seems to work well enough for me. But if I have hundreds of tabs open in Firefox (sometimes I do this) I can get to the point where Firefox eats about 5 GB+ of memory. I will start maxing out physical memory and will start eating swap. Somewhere around where I start eating about 50% of swap things really start slowing down and any click in Firefox takes a couple seconds. At least 2 of the CPU cores (out of 4) will get maxed out. Sometimes all will get maxed out briefly. Once I restart Firefox the system performs well again and even empties out the swap relatively quickly.

I tried operating without a swap space but this only resulted in frequent OOM conditions which sometimes took minutes before the OOM killer was actually invoked. Until that time all four cores would often get maxed at 100% and the system would be unresponsive until the OOM finally kicked in (or I used SYSREQ)

I tried once to set vm.swappiness=10 or vm.swappiness=1 but this still resulted in OOM conditions oddly enough when there was still plenty of free swap.

I set vm.min_free_kbytes=262144 very high like this to make the system much more resistant to OOM conditions. I no longer have problems with OOM conditions although I do have problems with responsiveness as indicated above.

Any way that is my story. 😊 If you have any advice please let me know or if you want to share your setup and settings and how they work. I would love to be able to get the system to perform better (i.e. where I'm not eating CPU) even when I am at 95% RAM use and 60% swap. I suspect to a certain extent I just need more physical ram and I'm just asking too much of what I have?

Last edited by davidm (2014-07-23 01:41:11)

Why I support GNU/Linux

Offline

#2 2014-07-23 02:03:09

lucke

Member

From: Poland

Registered: 2004-11-30

Posts: 4,018

Alt-SysRq-f calls an oom killer.

See if booting with zswap.enabled=1 in the kernel line helps.

Use dstat (e.g. "dstat -cdnpgms --top-bio --top-cpu --top-mem") to see what exactly is happening when your CPU cores get maxed out and what is maxing them out.

Last edited by lucke (2014-07-23 02:03:44)

Offline

#3 2014-07-23 02:12:56

davidm

Member

Registered: 2009-04-25

Posts: 366

Thanks for the reply. I don't have any OOM problems now at the moment. Only the nasty CPU usage when maxing physical RAM and > 50-60% swap. The dstat command is a great tip. I'll do that the next time I trigger it. It might be related to btrfs and my loopback swap file perhaps?

From researching I found 'vm.vfs_cache_pressure' isn't exactly what I thought it was.

SUSE has some good documentation for this (for the benefit of anyone searching for info) :

15.3.1. Reclaim Ratios¶
/proc/sys/vm/swappiness

This control is used to define how aggressively the kernel swaps out anonymous memory relative to pagecache and other caches. Increasing the value increases the amount of swapping. The default value is 60.

Swap I/O tends to be much less efficient than other I/O. However, some pagecache pages will be accessed much more frequently than less used anonymous memory. The right balance should be found here.

If swap activity is observed during slowdowns, it may be worth reducing this parameter. If there is a lot of I/O activity and the amount of pagecache in the system is rather small, or if there are large dormant applications running, increasing this value might improve performance.

Note that the more data is swapped out, the longer the system will take to swap data back in when it is needed.

/proc/sys/vm/vfs_cache_pressure

This variable controls the tendency of the kernel to reclaim the memory which is used for caching of VFS caches, versus pagecache and swap. Increasing this value increases the rate at which VFS caches are reclaimed.

It is difficult to know when this should be changed, other than by experimentation. The slabtop command (part of the package procs) shows top memory objects used by the kernel. The vfs caches are the "dentry" and the "*"inode_cache" objects. If these are consuming a large amount of memory in relation to pagecache, it may be worth trying to increase pressure. Could also help to reduce swapping. The default value is 100.

/proc/sys/vm/min_free_kbytes

This controls the amount of memory that is kept free for use by special reserves including “atomic” allocations (those which cannot wait for reclaim). This should not normally be lowered unless the system is being very carefully tuned for memory usage (normally useful for embedded rather than server applications). If “page allocation failure” messages and stack traces are frequently seen in logs, min_free_kbytes could be increased until the errors disappear. There is no need for concern, if these messages are very infrequent. The default value depends on the amount of RAM.

[http://doc.opensuse.org/products/draft/ ... vm.reclaim](http://doc.opensuse.org/products/draft/...vm.reclaim)

I wonder too if I might have went overboard with setting 'proc/sys/vm/vfs_cache_pressure=400' and this is causing too much activity as things fill? I may need to experiment I guess.

Why I support GNU/Linux

Offline

#4 2014-07-23 02:32:23

lucke

Member

From: Poland

Registered: 2004-11-30

Posts: 4,018

I don't think playing with these knobs could offer you much, and it's probably best to keep them at their defaults. You (probably) simply want to use more memory than you have. vfs_cache_pressure shouldn't do much for you, I am not sure if increasing min_free_kbytes improves the situation in regards to oom, it might make things worse instead (it's needed to make sure the system has enough memory to do certain things, if you set it higher, then you have effectively less memory). With high swappiness it swaps out unused parts of processes earlier, that means you run out of memory later. You normally shouldn't expect to be able to put a lot of a process in swap and be able to work with it, and this is apparently what happens with your Firefox. zswap might help, because, in a way, it gives you more memory.

dstat can help you figure out what is happening.

Also have a look at smem ("smem -kt"), it can show you nicely what is in your swap.

Last edited by lucke (2014-07-23 21:57:26)

Offline

#5 2014-07-23 20:22:36

davidm

Member

Registered: 2009-04-25

Posts: 366

I went ahead and put swappiness and vfs_cache_pressure back to their defaults. I decided to adjust min_free_kbytes to about 6% of my memory so it's set on 320000 (320MB). I've seen this recommended as a guideline although by default I suspect it is set much lower. The 320 MB of memory really will not kill me and the memory is still technically usable it's just that the kernel attempts to keep that much free. I believe it was only when I raised this number initially to 'vm.min_free_kbytes=262144' that OOM issues went away.

One thing is that I use btrfs and LZO compression. I have yet to turn off CoW for the swap file so perhaps I should do that. But also I suspect that due to the way it works it might use a little more memory and CPU for I/O operations.

This would be a neat thing to do some systematic tests with sometime.

Why I support GNU/Linux

Offline

#6 2014-07-24 21:12:04

firekage

Member

From: Eastern Europe, Poland
Registered: 2013-06-30
Posts: 543

Sorry for posting but i would like to ask for one thing. Where did you put config file for these:

```
vm.swappiness=50
vm.vfs_cache_pressure=400
vm.min_free_kbytes=262144
```

I know that, for example, vm.swappiness can be set by typing `sudo systemctl vm.swappiness=50` in terminal, but where are files for it? I would like to put in in order to have it after reboot.

Thanks.
Offline

#7 2014-07-24 22:26:16

lucke
Member
From: Poland
Registered: 2004-11-30
Posts: 4,018

You can put it in `/etc/sysctl.d/99-sysctl.conf`. <https://wiki.archlinux.org/index.php/Sysctl>

Offline

#8 2014-07-31 15:24:55

ooo
Member
Registered: 2013-04-10
Posts: 1,607

vfs_cache_pressure value larger than 100 may negative performance impact: <https://www.kernel.org/doc/Documentation/sysctl/vm.txt>

```
vfs_cache_pressure
-----
```

This percentage value controls the tendency of the kernel to reclaim the memory which is used for caching of directory and inode objects.

At the default value of `vfs_cache_pressure=100` the kernel will attempt to reclaim dentries and inodes at a "fair" rate with respect to pagecache and swapcache reclaim. Decreasing `vfs_cache_pressure` causes the kernel to prefer to retain dentry and inode caches. When `vfs_cache_pressure=0`, the kernel will never reclaim dentries and inodes due to memory pressure and this can easily lead to out-of-memory conditions. Increasing `vfs_cache_pressure` beyond 100 causes the kernel to prefer to reclaim dentries and inodes.

Increasing `vfs_cache_pressure` significantly beyond 100 may have negative performance impact. Reclaim code needs to take various locks to find freeable directory and inode objects. With `vfs_cache_pressure=1000`, it will look for ten times more freeable objects than there are.

`vfs_cache_pressure=50` is recommended here with a simple test case: <http://rudd-o.com/linux-and-free-softwa...o-fix-that>

In either case the optimal value depends on your system and workloads. Generally you should only tweak `vfs_cache_pressure` if you have performance issues related to disk caching.

I agree with lucke that enabling [zswap](#) could help when you're getting low on ram and just can't close all those firefox tabs.

Offline
Pages: 1

- [Index](#)
- » [System Administration](#)
- » [tuning vm.swappiness; vm.vfs_cache_pressure; vm.min_free_kbytes](#)

Jump to

System Administration

▼

Go

[Atom topic feed](#)

Powered by [FluxBB](#)