

Table Of Contents

Table Of Contents ..... 1

Say Hello To /proc/sys/vm ..... 2

    pdflush ..... 2

    swappiness ..... 3

    dirty\_ratio ..... 3

    Making Changes To VM Permanently ..... 3

    References: ..... 3

[Home](#) > [Faq](#) > [BASH Shell](#)

## Linux Tuning The VM (memory) Subsystem

Posted by [Vivek Gite](#) <[vivek@nixcraft.com](mailto:vivek@nixcraft.com)>

I've fast RAID-10 disk subsystem with multiple SCSI disks. Apps running under modern Linux kernel don't write directly to the disk. They write it to the file system cache which is managed by Linux kernel virtual memory manager. Since I've high performance RAID controller I need to decrease the number of flushes. How do I tune virtual memory subsystem under Linux operating systems for better performance?



[1]

Linux allows you to tune the VM subsystem. However, tuning the memory subsystem is a challenging task. Wrong settings can affect the overall performance of your system. I suggest you modify one setting at a time and monitor your system for sometime. If performance increased keep the settings else revert back.

## Say Hello To /proc/sys/vm

The files in this directory can be used to tune the operation of the virtual memory (VM) subsystem of the Linux kernel:

```
cd /proc/sys/vm
ls -l
```

Sample outputs:

```
total 0
-rw-r--r-- 1 root root 0 Oct 16 04:21 block_dump
-rw-r--r-- 1 root root 0 Oct 16 04:21 dirty_background_ratio
-rw-r--r-- 1 root root 0 Oct 16 04:21 dirty_expire_centisecs
-rw-r--r-- 1 root root 0 Oct 16 04:21 dirty_ratio
-rw-r--r-- 1 root root 0 Oct 16 04:21 dirty_writeback_centisecs
-rw-r--r-- 1 root root 0 Oct 16 04:21 drop_caches
-rw-r--r-- 1 root root 0 Oct 16 04:21 flush_mmap_pages
-rw-r--r-- 1 root root 0 Oct 16 04:21 hugetlb_shm_group
-rw-r--r-- 1 root root 0 Oct 16 04:21 laptop_mode
-rw-r--r-- 1 root root 0 Oct 16 04:21 legacy_va_layout
-rw-r--r-- 1 root root 0 Oct 16 04:21 lowmem_reserve_ratio
-rw-r--r-- 1 root root 0 Oct 16 04:21 max_map_count
-rw-r--r-- 1 root root 0 Oct 16 04:21 max_writeback_pages
-rw-r--r-- 1 root root 0 Oct 16 04:21 min_free_kbytes
-rw-r--r-- 1 root root 0 Oct 16 04:21 min_slab_ratio
-rw-r--r-- 1 root root 0 Oct 16 04:21 min_unmapped_ratio
-rw-r--r-- 1 root root 0 Oct 16 04:21 mmap_min_addr
-rw-r--r-- 1 root root 0 Oct 16 04:21 nr_hugepages
-r--r--r-- 1 root root 0 Oct 16 04:21 nr_pdflush_threads
-rw-r--r-- 1 root root 0 Oct 16 04:21 overcommit_memory
-rw-r--r-- 1 root root 0 Oct 16 04:21 overcommit_ratio
-rw-r--r-- 1 root root 0 Oct 16 04:21 pagecache
-rw-r--r-- 1 root root 0 Oct 16 04:21 page-cluster
-rw-r--r-- 1 root root 0 Oct 16 04:21 panic_on_oom
-rw-r--r-- 1 root root 0 Oct 16 04:21 percpu_pagelist_fraction
-rw-r--r-- 1 root root 0 Oct 16 04:21 swappiness
-rw-r--r-- 1 root root 0 Oct 16 04:21 swap_token_timeout
-rw-r--r-- 1 root root 0 Oct 16 04:21 vfs_cache_pressure
-rw-r--r-- 1 root root 0 Oct 16 04:21 zone_reclaim_mode
```

## pdflush

Type the following command to see current wake up time of pdflush:

```
# sysctl vm.dirty_background_ratio
```

Sample outputs:

```
sysctl vm.dirty_background_ratio = 10
```

vm.dirty\_background\_ratio contains 10, which is a percentage of total system memory, the number of pages at which the pdflush background writeback daemon will start writing out dirty data. However, for fast RAID based disk system this may cause large flushes of dirty memory pages. If you increase this value from 10 to 20 (a large value) will result into less frequent flushes:

```
# sysctl -w vm.dirty_background_ratio=20
```

## swappiness

Type the following command to see current default value:

```
# sysctl vm.swappiness
```

Sample outputs:

```
vm.swappiness = 60
```

The value 60 defines how aggressively memory pages are swapped to disk. If you do not want swapping, than lower this value. However, if your system process sleeps for a long time you may benefit with an aggressive swapping behavior by increasing this value. For example, you can change swappiness behavior by increasing or decreasing the value:

```
# sysctl -w vm.swappiness=100
```

## dirty\_ratio

Type the following command:

```
# sysctl vm.dirty_ratio
```

Sample outputs:

```
vm.dirty_ratio = 40
```

The value 40 is a percentage of total system memory, the number of pages at which a process which is generating disk writes will itself start writing out dirty data. This is nothing but the ratio at which dirty pages created by application disk writes will be flushed out to disk. A value of 40 mean that data will be written into system memory until the file system cache has a size of 40% of the server's RAM. So if you've 12GB ram, data will be written into system memory until the file system cache has a size of 4.8G. You change the dirty ratio as follows:

```
# sysctl -w vm.dirty_ratio=25
```

## Making Changes To VM Permanently

You need to add the settings to /etc/sysctl.conf. See our previous FAQ [making changes to /proc filesystem](#) <sup>[2]</sup> permanently.

### References:

1. [The /proc filesystem](#)
2. man page sysctl

4000+ howtos and counting! Want to read more Linux / UNIX howtos, tips and tricks? Subscribe to our [daily email](#) newsletter or [weekly newsletter](#) to make sure you don't miss a single tip/tricks. Alternatively, subscribe via [RSS/XML](#) feed.

Article printed from Frequently Asked Questions About Linux / UNIX: <http://www.cyberciti.biz/faq/>

URL to article: <http://www.cyberciti.biz/faq/linux-kernel-tuning-virtual-memory-subsystem/>

URLs in this post:

[1] Image: <http://www.cyberciti.biz/faq/category/linux/>

[2] making changes to /proc filesystem: <http://www.cyberciti.biz/faq/making-changes-to-proc-filesystem-permanently/>

---

Copyright © 2006-2010 [nixCraft](http://www.cyberciti.biz/). All rights reserved. This print / pdf version is for personal non-commercial use only. More details <http://www.cyberciti.biz/tips/copyright>.