

Linux Performance Tuning

Jaqui Lynch

USCMG

Jaqui.lynch@mainline.com

Agenda

- Introduction
- Various Commands
 - CPU
 - Memory
 - Disk and I/O
 - Network

Introduction

- Linux is Linux
 - Suse, Redhat, Debian, IBM
 - Security and performance
- Capacity & Scalability Viewpoint
 - XPZ
 - Test on X, Develop on P, Production on Z
- Much the same as tuning any other system
 - Processor, Memory, I/O, network
- Focus of this presentation is Redhat Fedora

Kernel Changes

- Redhat
 - /etc/sysctl.conf
- Echo "value" >/proc/kernel/parameter
 - Is reset at boot
- Remove all unneeded services
 - /etc/xinetd.d/
 - X, pop,

Commands for CPU

- vmstat
- top
- ulimit
- Provided by sysstat rpm
 - mpstat
 - iostat
 - sar
- /proc/cpuinfo
- ps

ulimit -a

ulimit -a

core file size	(blocks, -c)	0
data seg size	(kbytes, -d)	unlimited
file size	(blocks, -f)	unlimited
max locked memory	(kbytes, -l)	unlimited
max memory size	(kbytes, -m)	unlimited
open files	(-n)	1024
pipe size	(512 bytes, -p)	8
stack size	(kbytes, -s)	10240
cpu time	(seconds, -t)	unlimited
max user processes	(-u)	3072
virtual memory	(kbytes, -v)	unlimited

ulimit -u unlimited

Set max procs per user to unlimited

sar -v Kernel Values

Linux 2.4.20-30.9 (biteme.xxx.com) 03/30/2004

```
12:13:28 PM dentunusd  file-sz  %file-sz  inode-sz
12:13:30 PM      115509      107      0.27      118573
12:13:32 PM      115509      107      0.27      118573
12:13:34 PM      115509      107      0.27      118573
12:13:36 PM      115509      107      0.27      118573
12:13:38 PM      115509      107      0.27      118573
.....
Average:      115509      107      0.27      118573

12:13:28 PM super-sz  %super-sz  dquot-sz  %dquot-sz  rtsig-sz  %rtsig-sz
12:13:30 PM      0      0.00      0      0.00      0      0.00
12:13:32 PM      0      0.00      0      0.00      0      0.00
12:13:34 PM      0      0.00      0      0.00      0      0.00
12:13:36 PM      0      0.00      0      0.00      0      0.00
12:13:38 PM      0      0.00      0      0.00      0      0.00
.....
Average:      0      0.00      0      0.00      0      0.00
```

more /proc/cpuinfo

```
processor      : 0
vendor_id     : GenuineIntel
cpu family    : 6
model         : 7
model name    : Pentium III (Katmai)
cpu MHz       : 448.972
cache size    : 512 KB
```

Above are just a few of the items shown

CPU Time

- Real
 - Wallclock time
- User state CPU
 - Time spent running a user's program
 - Includes library calls
 - Can be affected by compile optimization and writing efficient code
- System code CPU
 - Time spent in the system state on the user's behalf
 - Kernel calls & all I/O routines
 - Can be affected by using blocking for I/O transfers
- I/O Time & Network Time
 - Time spent moving data and servicing I/O requests

uptime

- 12:23pm up 86 days, 4:42, 4 users, load average: 0.44, 0.61, 0.68
- 11:53am up 56 days 4:59 1 user load average: 5.45, 5.45, 5.46
- 09:20am up 23 days 15:23 1 user load average: 0.61, 0.65, 0.56

- Provides system load average
- Rough Indicator
- Counts as runnable all jobs waiting on disk or nfs I/O
- Last minute, last 5 minutes, last 15 minutes
 - Single user workstation <2
 - <3 Light 4-7 heavy >7 ??????

```

File Edit View Window Help
Quick Connect Profiles
[root@bitene rpms]# top
11:43:08 up 23:33, 2 users, load average: 0.47, 0.23, 0.15
62 processes: 60 sleeping, 2 running, 0 zombie, 0 stopped
CPU states: 100.0% user 0.0% system 0.0% nice 0.0% iowait 0.0% idle
Mem: 384568k av, 371960k used, 12608k free, 0k shrd, 90320k buff
Swap: 1020116k av, 1828k used, 1018288k free 128528k cached

  PID USER      PRI  NI  SIZE  RSS SHARE STAT  %CPU %MEM    TIME CPU  COMMAND
 13622 jaqui    25   0 5284 5284  236 R    99.5  1.3   0:32  0  phtestmae
 13624 root     16   0 1108 1108  868 R    0.3  0.2   0:00  0  top
    1 root     15   0  112   84   56 S    0.0  0.0   0:04  0  init
    2 root     15   0   0     0   0 SW    0.0  0.0   0:00  0  keventd
    3 root     15   0   0     0   0 SW    0.0  0.0   0:00  0  kapid
    4 root     34  19   0     0   0 SWM   0.0  0.0   0:00  0  ksoftirqd_CPU0
    9 root     25   0   0     0   0 SW    0.0  0.0   0:00  0  bdflush
    5 root     15   0   0     0   0 SW    0.0  0.0   0:00  0  kswapd
    6 root     15   0   0     0   0 SW    0.0  0.0   0:00  0  kscand/DMA
    7 root     15   0   0     0   0 SW    0.0  0.0   0:32  0  kscand/Normal
    8 root     15   0   0     0   0 SW    0.0  0.0   0:00  0  kscand/HighMem
   10 root     15   0   0     0   0 SW    0.0  0.0   0:00  0  kupdated
   11 root     25   0   0     0   0 SW    0.0  0.0   0:00  0  mdrecoveryd
   15 root     15   0   0     0   0 SW    0.0  0.0   0:00  0  kjournald
   73 root     24   0   0     0   0 SW    0.0  0.0   0:00  0  khubd
  214 root     15   0   0     0   0 SW    0.0  0.0   0:00  0  kjournald
  215 root     15   0   0     0   0 SW    0.0  0.0   0:00  0  kjournald
  216 root     15   0   0     0   0 SW    0.0  0.0   0:00  0  kjournald
  217 root     15   0   0     0   0 SW    0.0  0.0   0:00  0  kjournald
  218 root     15   0   0     0   0 SW    0.0  0.0   0:00  0  kjournald
[root@bitene rpms]#

```

Connected to 192.168.1.80 SSH2 - aes128-cbc - hmac-md5 - none 104x31 NUM

top on linux

```

12:24pm up 86 days, 4:43, 4 users, load average: 0.66, 0.63, 0.68
119 processes: 115 sleeping, 3 running, 0 zombie, 1 stopped
CPU states: 25.3% user, 7.6% system, 0.0% nice, 67.0% idle
Mem: 517356K av, 514272K used, 3084K free, 550780K shrd, 33728K buff
Swap: 525288K av, 6916K used, 518372K free 152520K cached

  PID USER      PRI  NI  SIZE  RSS SHARE STAT  LIB %CPU %MEM    TIME COMMAND
 22339 sybase    19   0 16384 15M 16236 R      0 12.8  3.1 3095m datasv
 14590 jaqui     4   0  1268 1268  1016 S      0  7.9  0.2   0:00 ssh
 14511 webid    12   0 12448 12M  8828 S      0  3.8  2.4   0:00 httpd
 25935 webid     7   0 14828 14M  6888 S      0  2.7  2.8   2:43 httpd
 14586 root      6   0  1060 1060   836 R      0  2.3  0.2   0:00 top
 10589 webid     2   0 13256 12M  7348 S      0  0.5  2.5   0:26 httpd
 14514 webid     1   0 12588 12M  8616 S      0  0.3  2.4   0:00 httpd
 14585 webid     0   0 11336 11M 11000 S      0  0.3  2.1   0:00 httpd
 25930 root      1   0 11160 10M 11008 S      0  0.1  2.1   0:11 httpd
 25938 webid     0   0 14740 14M  7072 S      0  0.1  2.8   2:31 httpd
 25970 webid     0   0 15548 15M  6868 S      0  0.1  3.0   2:16 httpd
 13675 root      0   0  1272 1240  1028 R      0  0.1  0.2   0:01 sshd

```

nmon

```

File Edit View Window Help
-----
Quick Connect Profiles
-----
nmon Linux9a Hostname=biteme Refresh=2.0secs 11:09:05
-----
CPU Utilisation
-----
CPU User% Sys% Wait% Idle|0 125 150 175 100|
0 0.0 0.0 0.0 100.0|>
-----
Memory State
-----
RAM High Low Swap
Total MB 376.1 0.0 376.1 996.2
Free MB 11.8 0.0 11.8 996.1
Free Percent 3.0% 0.0% 3.0% 99.0%
-----
shARED= 0.0 cached= 52.0 active= 38.4 bigfree= 0.0
buffers= 60.2 swpcached= 0.0 inactive= 101.4
-----
Network I/O
-----
I/F Name Recv=KB/s Trans=KB/s packin packout insize outsize Peak->Recv Trans
lo 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
eth0 0.0 0.1 0.5 0.5 60.0 182.0 7.4 7.3
-----
Filesystem and GPS
-----
Filesystem Size MB Free MB %Used %Inodes MountPoint
LABEL=/ 4918.3 4672.2 5.0% 4.4% /
LABEL=/boot 887.3 861.4 3.0% 0.1% /boot
none 0.0 0.0 nan% nan% /dev/pts
LABEL=/home 4918.2 4229.5 14.0% 0.1% /home
LABEL=/opt 980.5 964.4 1.6% 0.0% /opt
none 0.0 0.0 nan% nan% /proc
none 188.0 188.0 0.0% 0.0% /dev/shm
LABEL=/tmp 1968.3 1936.7 1.6% 0.1% /tmp
LABEL=/usr 4918.2 1652.3 66.4% 27.6% /usr
LABEL=/usr/local 9844.2 3805.7 61.3% 3.1% /usr/local
LABEL=/usr/local/softw 9844.2 7109.3 27.8% 0.4% /usr/local/software
LABEL=/var 1968.8 1658.2 15.8% 2.0% /var
swap not mounted
/dev/cdrom 4918.3 4672.2 5.0% 4.4% /mnt/cdrom
/dev/hdd4 4918.3 4672.2 5.0% 4.4% /mnt/zip
/dev/fd0 4918.3 4672.2 5.0% 4.4% /mnt/floppy
Disk I/O
-----
all data is Kbytes per second
DiskName Busy Read Write |0 125 150 175 100|
dev-3-0 0% 0.0 0.0KB| | | | |
dev-3-1 0% 0.0 0.0KB| | | | |
-----
Connected to 192.168.1.80 SSH2 - aes128-cbc - hmac-md5 - none 113x40

```

mpstat

mpstat 2 10 needs sysstat

mpstat 2 10 (phntest)

Linux 2.4.20-30.9 (biteme.xxx.com) 03/30/2004

	CPU	%user	%nice	%system	%idle	intr/s
01:33:28 PM	all	100.00	0.00	0.00	0.00	104.00
01:33:30 PM	all	100.00	0.00	0.00	0.00	100.00
01:33:32 PM	all	100.00	0.00	0.00	0.00	100.00
01:33:34 PM	all	100.00	0.00	0.00	0.00	102.50
01:33:36 PM	all	100.00	0.00	0.00	0.00	100.00
01:33:38 PM	all	100.00	0.00	0.00	0.00	100.00
01:33:40 PM	all	100.00	0.00	0.00	0.00	101.50
01:33:42 PM	all	100.00	0.00	0.00	0.00	100.00
01:33:44 PM	all	99.50	0.00	0.50	0.00	101.50
01:33:46 PM	all	100.00	0.00	0.00	0.00	100.00
01:33:48 PM	all	100.00	0.00	0.00	0.00	100.00
Average:	all	99.95	0.00	0.05	0.00	100.95

mpstat -P ALL

```
mpstat -P ALL 2 10          needs sysstat

Linux 2.2.17 (jaqui.abc.com)    05/15/01

11:29:18 AM CPU %user %nice %system %idle intr/s
11:29:20 AM all 1.50 0.00 3.25 95.25 102.00
11:29:20 AM 0 0.00 0.00 0.00 100.00 102.00
11:29:20 AM 1 3.00 0.00 6.50 90.50 102.00

11:29:20 AM CPU %user %nice %system %idle intr/s
11:29:22 AM all 1.50 0.00 1.00 97.50 101.50
11:29:22 AM 0 0.00 0.00 0.00 100.00 101.50
11:29:22 AM 1 3.00 0.00 2.00 95.00 101.50

.....
Average: CPU %user %nice %system %idle intr/s
Average: all 0.30 0.00 0.42 99.28 103.00
Average: 0 0.00 0.00 0.00 100.00 103.00
Average: 1 0.60 0.00 0.85 98.55 103.00
```

sar -u

Sar does not come with most Linux systems - you need to download it as part of the sysstat rpm

Create file using `sar -a -o sardata.out`

```
sar -u -f sardata.out
```

```
Linux 2.4.20-30.9 (biteme.xxx.com)    03/30/2004

12:34:44 CPU %user %nice %system %idle
12:34:46 all 53.50 0.00 9.00 37.50
12:34:48 all 79.00 0.00 11.50 9.50
12:34:50 all 80.00 0.00 6.50 13.50
12:34:52 all 59.00 0.00 8.00 33.00
12:34:54 all 75.50 0.00 15.50 9.00
Average: all 69.40 0.00 10.10 20.50
```

sar -q

```
# sar -q -f sardata.out
```

```
12:13:28 PM runq-sz plist-sz ldavg-1 ldavg-5
12:13:30 PM      1      63      1.03      0.96
12:13:32 PM      1      63      1.03      0.96
12:13:34 PM      1      63      1.03      0.96
12:13:36 PM      1      63      1.03      0.96
12:13:38 PM      1      63      1.02      0.96
.....
12:14:02 PM      1      63      1.02      0.96
12:14:04 PM      1      63      1.02      0.96
12:14:06 PM      1      63      1.02      0.96
12:14:08 PM      1      63      1.01      0.96
Average:         1      63      1.02      0.96
```

ps aux

```
ps aux
```

USER	PID	%CPU	%MEM	VSZ	RSS	TTY	STAT	START	TIME	COMMAND
root	1	0.0	0.0	1096	52	?	S	Feb18	0:03	init [3]
root	2	0.0	0.0	0	0	?	SW	Feb18	4:31	[kflushd]
root	3	0.0	0.0	0	0	?	SW	Feb18	0:00	[kpiod]
root	4	0.0	0.0	0	0	?	SW	Feb18	7:54	[kswapd]
root	5	0.0	0.0	0	0	?	SW<	Feb18	0:00	[mdrecoveryd]
bin	254	0.0	0.0	1092	216	?	S	Feb18	0:01	portmap
root	301	0.0	0.0	1280	372	?	S	Feb18	0:24	syslogd -m 0
jaqui	29044	0.0	0.1	1264	548	pts/3	S	10:08	0:00	-ksh
webid	29173	0.5	2.4	14432	12640	?	S	10:14	0:02	/usr/local/apache/bin/httpd -d /usr/local/apache
webid	29174	0.3	2.4	14840	12664	?	S	10:14	0:01	/usr/local/apache/bin/httpd -d /usr/local/apache
webid	29175	0.2	2.4	14728	12580	?	S	10:14	0:01	/usr/local/apache/bin/httpd -d /usr/local/apache
webid	29176	0.3	2.3	13908	12308	?	S	10:14	0:01	/usr/local/apache/bin/httpd -d /usr/local/apache
webid	29177	0.2	2.4	14724	12612	?	S	10:14	0:00	/usr/local/apache/bin/httpd -d /usr/local/apache
webid	29178	0.1	2.3	13784	12148	?	S	10:14	0:00	/usr/local/apache/bin/httpd -d /usr/local/apache
webid	29195	0.1	2.3	13888	12208	?	S	10:15	0:00	/usr/local/apache/bin/httpd -d /usr/local/apache
webid	29206	0.4	2.4	14728	12788	?	S	10:15	0:01	/usr/local/apache/bin/httpd -d /usr/local/apache

Commands for Memory

- /proc/meminfo
- sar
- /proc/sys/vm/*
- vmstat
- lsmod or look at /proc/modules
- vmstat
- swapon & swapoff
- free
- top

UNIX and Memory

- Uses ALWAYS FULL model
- Kernel Space
 - Memory reserved for the kernel
 - Used for device drivers, scheduling and to keep the kernel in memory
- User Space
 - Memory for application code – includes servers
- Buffer Space
 - Used to buffer I/O for disks and network
 - Used for DMA transfers between devices
 - Applied to both read and write operations

vmstat (Linux)

```
vmstat 2 10
```

Can use `vmstat -m` or `vmstat -g` to see in MB or GB

procs				memory			swap		io			system			cpu	
r	b	w	swpd	free	buff	cache	si	so	bi	bo	in	cs	us	sy	id	
1	0	0	6916	13700	27876	156100	0	0	4	4	1	4	3	4	1	
2	0	0	6916	9520	28196	160324	0	0	2121	348	202	138	79	4	17	
4	0	0	6916	4020	28560	163352	0	0	1520	0	219	272	80	15	5	
3	0	0	6916	1732	28532	165408	0	0	1226	250	234	208	73	8	20	
1	2	0	6916	3016	25628	164360	0	0	1250	569	247	395	78	9	13	
2	1	0	6916	2296	21976	164476	0	0	1430	485	245	250	79	11	10	
2	0	0	6916	2444	20208	164480	0	0	660	140	372	396	74	8	18	
2	0	0	6916	2360	19084	165736	0	0	1835	134	211	248	82	9	9	
1	0	0	6916	1192	17908	168180	0	0	2443	161	185	138	88	5	7	
2	0	0	6916	2824	16000	168544	0	0	2122	251	202	160	81	4	14	

Swap Devices

- Swap device is a disk partition
- Can be a logical volume with some effort
- Swap device in `/etc/fstab` defaults to active
- `swapon -s` list active swap devices

Filename	Type	Size	Used	Priority
/dev/hdb3	partition	1020116	1872	-1

swapon and free

Free

Test:

	total	used	free	shared	buffers	cached
Mem:	517356	515076	2280	465748	10708	179340
-/+ buffers/cache:		325028	192328			
Swap:	525288	6916	518372			

Biteme:

Mem:	384568	372860	11708	0	93332	130044
-/+ buffers/cache:		149484	235084			
Swap:	1020116	1872	1018244			

free -m

	total	used	free	shared	buffers	cached
Mem:	505	477	28	443	36	120
-/+ buffers/cache:		320	184			
Swap:	512	6	506			

/proc/meminfo

more /proc/meminfo

```
Mem: 529772544 436813824 92958720 552947712 32686080 62316544
Swap: 537894912 7081984 530812928
MemTotal: 517356 kB
MemFree: 90780 kB
MemShared: 539988 kB
Buffers: 31920 kB
Cached: 60856 kB
SwapTotal: 525288 kB
SwapFree: 518372 kB
```

/proc/sys/vm

```
/proc/sys/vm
```

```
cat bdflush
```

```
50      500      0      0      500      3000      80      50      0
```

```
cat max-readahead
```

```
31
```

```
cat min-readahead
```

```
3
```

sar -B

```
sar -B -f sardata.out
```

12:13:28 PM	pgpgin/s	pgpgout/s	activepg	inadtypg	inaclnpg	inatarpg
12:13:30 PM	0.00	40.00	47258	0	1596	11614
12:13:32 PM	0.00	8.00	47260	0	1596	11614
12:13:34 PM	0.00	8.00	47262	0	1596	11615
12:13:36 PM	0.00	16.00	47264	0	1596	11615
12:13:38 PM	0.00	8.00	47266	0	1596	11616
.....						
12:14:06 PM	0.00	8.00	47294	0	1596	11622
12:14:08 PM	0.00	8.00	47296	0	1596	11622
Average:	0.00	12.50	47277	0	1596	11618

```
sar -B -f sardata.out
```

12:34:44	pgpgin/s	pgpgout/s	activepg	inadtypg	inaclnpg	inatarpg
12:34:46	376.00	1.50	0	0	0	0
12:34:48	441.50	1.50	0	0	0	0
12:34:50	568.00	71.50	0	0	0	0
12:34:52	437.00	1.50	0	0	0	0
12:34:54	450.50	22.50	0	0	0	0
Average:	454.60	19.70	0	0	0	0

sar -r

```
sar -r -f sardata.out
```

12:13:28 PM	kbmem free	kbmem used	%mem used	kbmem shrd	kb buffers	kb cached	kbswp free	kbswp used	%swp used
12:13:30 PM	10268	374300	97.33	0	90408	129280	1018268	1848	0.18
12:13:32 PM	10268	374300	97.33	0	90416	129280	1018268	1848	0.18
12:13:34 PM	10268	374300	97.33	0	90424	129280	1018268	1848	0.18
12:13:36 PM	10268	374300	97.33	0	90432	129284	1018268	1848	0.18
12:13:38 PM	10268	374300	97.33	0	90440	129284	1018268	1848	0.18
.....									
12:14:06 PM	10268	374300	97.33	0	90552	129296	1018268	1848	0.18
12:14:08 PM	10268	374300	97.33	0	90560	129296	1018268	1848	0.18
Average:	10268	374300	97.33	0	90484	129288	1018268	1848	0.18

Commands for I/O

- iostat
- sar
- /proc/sys/fs/*
- /proc/partitions
- hdparm
- /proc/stat
- df

df

Df or df -k

Filesystem	1k-blocks	Used	Available	Use%	Mounted on
/dev/sda1	127880	52265	69011	43%	/
/dev/sda8	12762679	9957815	2141620	82%	/home
/dev/sda5	2971111	2097239	720222	74%	/usr
/dev/sda6	762122	277769	444981	38%	/var

df -h

Filesystem	Size	Used	Available	Use%	Mounted on
/dev/sda5	539M	155M	383M	29%	/
/dev/sdb6	21G	14G	7.5G	65%	/home
/dev/sda6	2.9G	1007M	1.9G	34%	/usr
/dev/sda9	1.9G	358M	1.6G	18%	/usr/local
/dev/sda8	992M	96M	896M	10%	/var
/dev/sdb7	19G	231M	19G	1%	/web

df -i Inodes

[root@biteme root]# df -i

Filesystem	Inodes	IUsed	IFree	IUse%	Mounted on
/dev/hdb2	640000	28460	611540	5%	/
/dev/hdb1	115712	64	115648	1%	/boot
/dev/hdb7	640000	769	639231	1%	/home
/dev/hdb11	127744	11	127733	1%	/opt
none	48140	1	48139	1%	/dev/shm
/dev/hdb10	256512	150	256362	1%	/tmp
/dev/hdb8	640000	176782	463218	28%	/usr
/dev/hdb5	1281696	43068	1238628	4%	/usr/local
/dev/hdb6	1281696	5634	1276062	1%	/usr/local/software
/dev/hdb9	256512	5262	251250	3%	/var

iostat

iostat 2 10

Linux 2.4.20-30.9 (biteme.wideopenwest.com) 03/30/2004

avg-cpu:	%user	%nice	%sys	%idle
	3.14	0.31	0.16	96.39

Device:	tps	Blk_read/s	Blk_wrtn/s	Blk_read	Blk_wrtn
dev3-0	0.00	0.00	0.00	256	0
dev3-1	0.82	4.90	9.69	447714	886136

avg-cpu:	%user	%nice	%sys	%idle
	100.00	0.00	0.00	0.00

Device:	tps	Blk_read/s	Blk_wrtn/s	Blk_read	Blk_wrtn
dev3-0	0.00	0.00	0.00	0	0
dev3-1	0.00	0.00	0.00	0	0

sar I/O rates

sar -b -f sardata.out

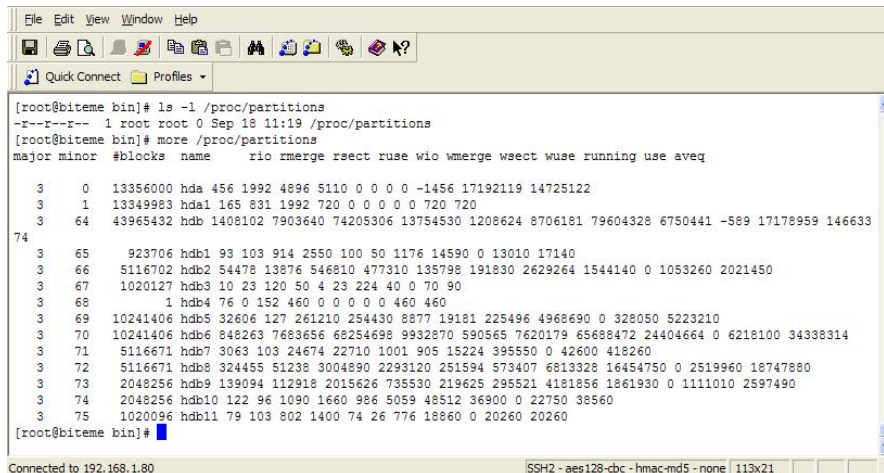
Linux 2.4.20-30.9 (biteme.xxx.com) 03/30/2004

	tps	rtps	wtps	bread/s	bwrtn/s
12:13:28 PM					
12:13:30 PM	3.50	0.00	3.50	0.00	80.00
12:13:32 PM	1.50	0.00	1.50	0.00	16.00
12:13:34 PM	1.50	0.00	1.50	0.00	16.00
12:13:36 PM	1.50	0.00	1.50	0.00	32.00
12:13:38 PM	1.50	0.00	1.50	0.00	16.00
12:13:40 PM	1.50	0.00	1.50	0.00	16.00
.....					
12:14:08 PM	1.50	0.00	1.50	0.00	16.00
Average:	1.88	0.00	1.88	0.00	25.00

/proc/sys/fs

- Max values can be set in /proc/sys/fs/*-max
- cat file-max
 - 39321
- cat file-nr
 - 425 114 39321
- cat inode-nr
 - 172688 72170
- hdparm
 - Use man hdparm
 - Sets filesystem read ahead count, drive write cache, drive lookahead prefetch count, drive lookahead

/proc/partitions



```
[root@biteme bin]# ls -l /proc/partitions
-r--r--r-- 1 root root 0 Sep 18 11:19 /proc/partitions
[root@biteme bin]# more /proc/partitions
major minor #blocks name          rio rmerge rsect ruse wio wmerge wsect wuse running use aveq
3        0    13356000 hda  456 1992 4896 5110 0 0 0 0 -1456 17192119 14725122
3        1    13349983 hda1 165 831 1992 720 0 0 0 0 0 0 720 720
3        64    43965432 hdb  1408102 7903640 74205306 13754530 1208624 8706181 79604328 6750441 -589 17178959 146633
74
3        65         923706 hdb1  93 103 914 2550 100 50 1176 14590 0 13010 17140
3        66    5116702 hdb2 54478 13876 546810 477310 135798 191830 2629264 1544140 0 1053260 2021450
3        67    1020127 hdb3  10 23 120 50 4 23 224 40 0 70 90
3        68         1 hdb4  76 0 152 460 0 0 0 0 0 460 460
3        69    10241406 hdb5 32606 127 261210 254430 8877 19181 225496 4968690 0 328050 5223210
3        70    10241406 hdb6 848263 7683656 68254698 9932870 590565 7620179 65688472 24404664 0 6218100 34338314
3        71    5116671 hdb7  3063 103 24674 22710 1001 905 15224 395550 0 42600 418260
3        72    5116671 hdb8 324455 51238 3004890 2293120 251594 573407 6813328 16454750 0 2519960 18747880
3        73    2048256 hdb9 139094 112918 2015626 735530 219625 295521 4181856 1861930 0 1111010 2597490
3        74    2048256 hdb10 122 96 1090 1660 986 5059 48512 36900 0 22750 38560
3        75    1020096 hdb11 79 103 802 1400 74 26 776 18860 0 20260 20260
[root@biteme bin]#
```

Isof

- Isof is a 3rd party tool that you can use to look at open files and network connections
- <ftp://vic.cc.purdue.edu/pub/tools/unix/Isof/README>
- `cd /usr/local`
- `tar -xvf soft/tarfiles/Isof_4.71.tar`
- `cd Isof_4.71`
- `./Configure linux`
- ensure you set HASSECURITY to yes
- `make`
- `cp Isof /usr/local/bin`
- `cp Isof.8 /usr/local/man/man8`
- Try `Isof | grep TCP` or `Isof | grep UDP` or `Isof | more`

Intra and Inter Policies

- Intra Policy
 - How data is laid out on the disk
 - Outer edge, outer middle, center, inner middle, inner edge
- Inter Policy
 - How data is laid out between/across disks

RAID Levels

- Raid-0
 - Disks combined into single volume stripeset
 - Data striped across the disks
- Raid-1
 - Every disk mirrored to another
 - Full redundancy of data but needs extra disks
 - At least 2 I/Os per random write
- Raid-0+1
 - Striped mirroring
 - Combines redundancy and performance

RAID Levels

- RAID-5
 - Data striped across a set of disks
 - 1 more disk used for parity bits
 - Parity may be striped across the disks also
 - At least 4 I/Os per random write (read/write to data and read/write to parity)
 - Uses hot spare technology

Other I/O performance notes

- Mirroring of disks
- Write verify
- SCSI (1 disk at a time) vs scsi-2 fw diff
- RAID and SSA
- I/O Pacing and Async I/O
- Striping

Networking

- netstat
- nfsstat
- lsof | grep TCP
- lsof | grep UDP

/proc/sys/net/core

- Critical on high speed networks such as z/390 hypersockets (24gb/sec) or gigabit ethernet
- Improves TCP throughput
- TCP window sizes (2.1.90 and >)
 - rmem_max & wmem_max
 - Default is 64kb – can increase it
- Default Buffer Settings
 - Max socket buffer size 64kb
 - Default TCP socket buffer size 32kb
 - Default UDP socket buffer size 32kb?

/proc/sys/net/core on Fedora

```
/proc/sys/net/core  
  
[jaqui@biteme core]$ cat optmem_max  
10240  
[jaqui@biteme core]$ cat rmem_default  
65535  
[jaqui@biteme core]$ cat rmem_max  
65535  
[jaqui@biteme core]$ cat wmem_default  
65535  
[jaqui@biteme core]$ cat wmem_max  
65535
```

/proc/sys/net/ipv4 on Fedora

```
/proc/sys/net/ipv4

[jaqui@biteme ipv4]$ cat ip_forward
0
[jaqui@biteme ipv4]$ cat tcp_rmem
4096      87380    174760
[jaqui@biteme ipv4]$ cat tcp_mem
97280    97792    98304
[jaqui@biteme ipv4]$ cat tcp_wmem
4096     16384    131072
[jaqui@biteme ipv4]$ cat tcp_window_scaling
1
```

netstat -i

Kernel Interface table

Iface	MTU	Met	RX-OK	RX-ERR	RX-DRP	RX-OVR	TX-OK	TX-ERR	TX-DRP	TX-OVR	Flg
eth0	1500	0	94408545	0	0	0	99872636	0	0	0	BRU
lo	3924	0	287611603	0	0	0	287611603	0	0	0	LRU

netstat -i (Fedora)

Kernel Interface table

Iface	MTU	Met	RX-OK	RX-ERR	RX-DRP	RX-OVR	TX-OK	TX-ERR	TX-DRP	TX-OVR	Flg
eth0	1500	0	12023	0	0	0	10340	0	0	0	BMRU
lo	16436	0	70	0	0	0	70	0	0	0	LRU

nfsstat

- Client and Server NFS Info
- `nfsstat -cn, -r, -s`
 - retransmissions due to errors
 - badcalls
 - timeouts
 - waits
 - reads

Useful References

- <http://www.tunelinux.com>
- <http://linuxperf.nl/linux.org>
- http://www.psc.edu/networking/perf_tune.html
 - Enabling high performance data transfers on hosts
- <http://linuxperf.nl/linux.org/links.html>
- <http://www-124.ibm.com/developerworks/opensource/linuxperf/>
- http://www.ibm.com/developerworks/eserver/articles/analyze_aix/

Summary

- Linux has useful free performance tools
- Download iptraf and sysstat toolsets

- jaqui@circle4.com
- <http://www.circle4.com/jaqui/>
- <http://www.mainline.com>