



# Inspecting a multi-everything Linux system

Aurimas Mikalauskas  
Zabbix Conference  
September 12-13, 2014

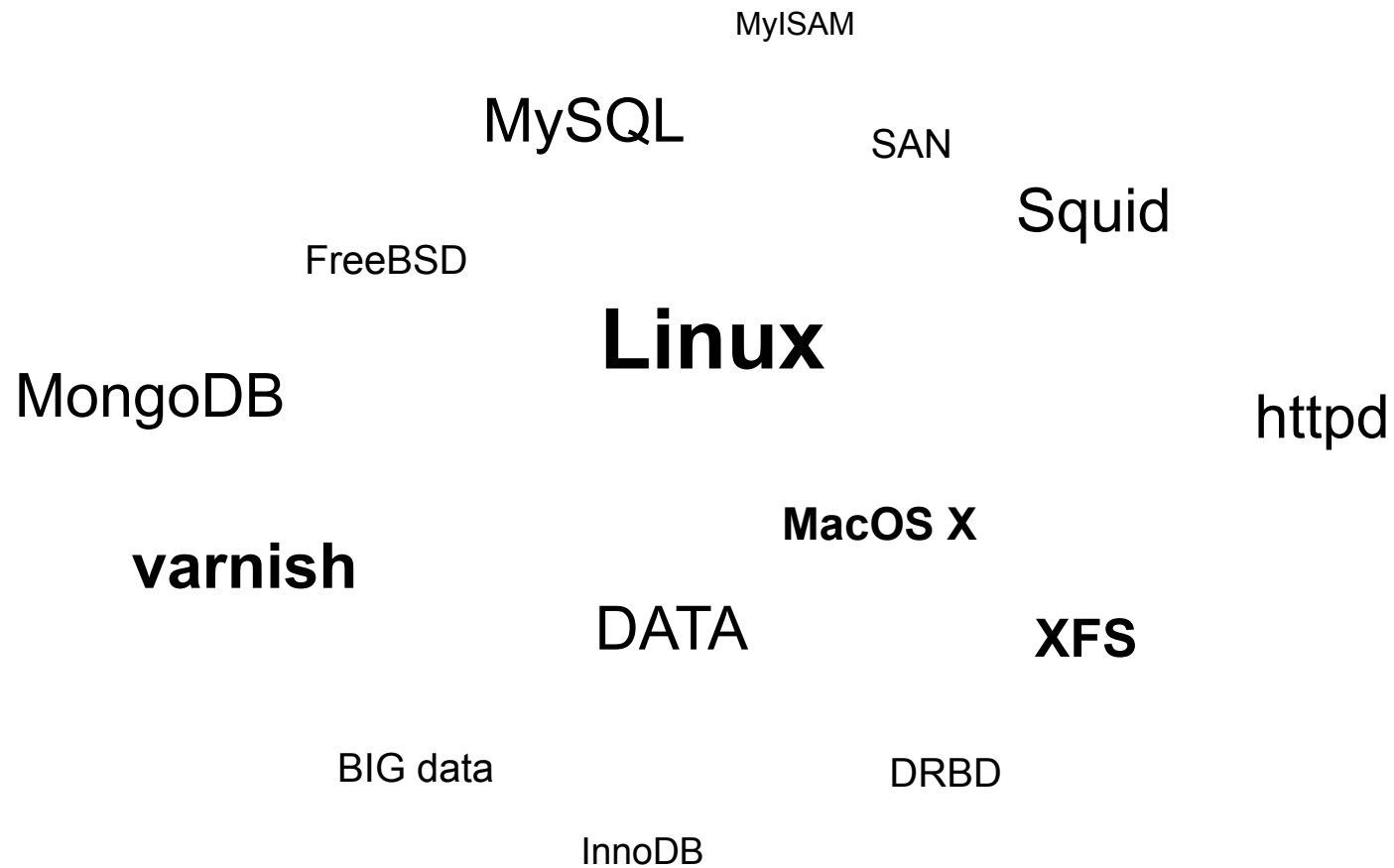
3

# LOAD AVERAGE

**LOAD AVERAGE**

**DISK I/O**

**IT'S JUST**  
**a system.**



**load average: 1.00, 1.00, 1.00**

# Quick Quiz

There's a server:

- It has 2 6-core CPUs (i.e. 12 cores)
- Hyper-threading is off (i.e. 12 cpu threads)
- 32GB of RAM
- 8 disks in RAID10



# Quick Quiz

There's a server:

- It has 2 6-core CPUs (i.e. 12 cores)
- Hyper-threading is off (i.e. 12 cpu threads)
- 32GB of RAM
- 8 disks in RAID10

Which load average value indicates server is **saturated**?

- a) 4.0
- b) 8.0
- c) 12.0
- d) 16.0
- e) 20.0

# Answer

Either of the them:

- a) 4.0
- b) 8.0
- c) 12.0
- d) 16.0
- e) 20.0

# Answer

Either of the them:

- a) 4.0            - 4 processes writing to disks
- b) 8.0
- c) 12.0
- d) 16.0
- e) 20.0

# Answer

Either of the them:

- a) 4.0            - 4 processes writing to disks
- b) 8.0            - 8 processes reading from disks
- c) 12.0
- d) 16.0
- e) 20.0

# Answer

Either of the them:

- a) 4.0            - 4 processes writing to disks
- b) 8.0            - 8 processes reading from disks
- c) 12.0           - 12 processes using CPU
- d) 16.0
- e) 20.0

# Answer

Either of the them:

- a) 4.0            - 4 processes writing to disks
- b) 8.0            - 8 processes reading from disks
- c) 12.0           - 12 processes using CPU
- d) 16.0           - 12 on CPU & 4 writing to disks
- e) 20.0

# Answer

Either of the them:

- a) 4.0            - 4 processes writing to disks
- b) 8.0            - 8 processes reading from disks
- c) 12.0           - 12 processes using CPU
- d) 16.0           - 12 on CPU & 4 writing to disks
- e) 20.0           - 12 on CPU & 8 reading from disks





# STANDARD I/O

library

what about

**UNBUFFERED I/O?**

# Writing big file with dd

```
# dd if=/dev/fioa of=/data/opt/testfile.bin bs=4096 count=7864320  
...  
32212254720 bytes (32 GB) copied, 70.7952 s, 455 MB/s
```

# Unbuffered I/O in real life

```
# vmstat 1
```

```
procs
r  b      swpd      free      buff      cache      si      so      bi      bo      in      cs      us      sy      id      wa      st
0  0         0 64979240 6457316  48724      0      0         0      0      726     177      0      0     100      0      0
1  0         0 64736104 6578076 168964      0      0    121456      0     1769    5468      0      2     98      0      0
1  0         0 63733356 7065556  656456      0      0    487424      0     6133   21445      0      7     93      1      0
1  0         0 62660484 7586900 1177756      0      0    521344     12     6051   22913      0      8     92      0      0
1  0         0 61592424 8105460 1696316      0      0    518656      0     6082   23024      0      7     93      0      0
1  0         0 60440380 8665556 2256852      0      0    559932     44     6266   26413      0      7     92      1      0
1  0         0 59119376 9306964 2898172      0      0    641348      0     7015   31229      0      6     93      0      0
1  0         0 58010208 9845716 3436984      0      0    538752      0     6158   25078      0      7     93      0      0
1  0         0 57018712 10327256 3918336      0      0    481536     24     6004   21398      0      7     93      0      0
1  0         0 55895652 10872924 4464236      0      0    545664      0     6280   25376      0      6     93      0      0
3  0         0 54591312 11506264 5097260      0      0    633344      8     6985   31139      0      8     90      1      0
2  0         0 53263628 12151256 5742488      0      0    644480      0     6971   31668      0      6     93      1      0
1  0         0 51964508 12782296 6373724      0      0    631424      0     7030   31187      0      8     91      1      0
1  0         0 51054968 13224664 6815484      0      0    442212      0     7619   27035      0      5     93      2      0
0  1         0 50891056 13304412 6895208      0      0     79644     12     8448   18104      0      2     92      6      0
0  1         0 49882144 13794140 7385144      0      0    489984   229598     9420   31855      0      6     92      2      0
1  0         0 48985468 14228012 7819192      0      0    433920   434756   10934   30660      0      5     93      1      0
1  1         0 48024344 14694492 8285256      0      0    466432   471500   10914   33897      0     12     84      3      0
1  0         0 47064056 15160924 8751692      0      0    466432   467400   10979   33871      0      7     90      2      0
... many samples skipped ...
r  b      swpd      free      buff      cache      si      so      bi      bo      in      cs      us      sy      id      wa      st
1  0         0  1940860 37169660 30761252      0      0   455680 459200   14158  27724      0      6     91      2      0
0  1         0   920864 37665532 31257060      0      0   495744 500200   14708  28307      0      6     92      2      0
2  0         0  3304476 35016696 31506636      0      0   248832 250136   10781  14505      0      9     89      1      0
1  0         0  9349772 28971956 31506812      0      0         0      0     3503     294      0      6     94      0      0
4  0         0 14827068 23494436 31506704      0      0         0      0    11120     181      0     14     86      0      0
```

# Reads kick in. Writes?

```
# vmstat 1
```

```
procs
r  b      swpd      free     buff     cache    si    so      bi     bo      in     cs    us    sy    id    wa    st
0  0         0 64979240 6457316  48724     0     0         0      0     726    177   0     0  100   0     0
1  0         0 64736104 6578076 168964     0     0    121456     0    1769   5468   0     2   98   0     0
1  0         0 63733356 7065556  656456     0     0    487424     0    6133  21445   0     7   93   1     0
1  0         0 62660484 7586900 1177756     0     0    521344    12    6051  22913   0     8   92   0     0
1  0         0 61592424 8105460 1696316     0     0    518656     0    6082  23024   0     7   93   0     0
1  0         0 60440380 8665556 2256852     0     0    559932    44    6266  26413   0     7   92   1     0
1  0         0 59119376 9306964 2898172     0     0    641348     0    7015  31229   0     6   93   0     0
1  0         0 58010208 9845716 3436984     0     0    538752     0    6158  25078   0     7   93   0     0
1  0         0 57018712 10327256 3918336     0     0    481536    24    6004  21398   0     7   93   0     0
1  0         0 55895652 10872924 4464236     0     0    545664     0    6280  25376   0     6   93   0     0
3  0         0 54591312 11506264 5097260     0     0    633344     8    6985  31139   0     8   90   1     0
2  0         0 53263628 12151256 5742488     0     0    644480     0    6971  31668   0     6   93   1     0
1  0         0 51964508 12782296 6373724     0     0    631424     0    7030  31187   0     8   91   1     0
1  0         0 51054968 13224664 6815484     0     0    442212     0    7619  27035   0     5   93   2     0
0  1         0 50891056 13304412 6895208     0     0     79644    12    8448  18104   0     2   92   6     0
0  1         0 49882144 13794140 7385144     0     0    489984  229598    9420  31855   0     6   92   2     0
1  0         0 48985468 14228012 7819192     0     0    433920  434756   10934  30660   0     5   93   1     0
1  1         0 48024344 14694492 8285256     0     0    466432  471500   10914  33897   0    12   84   3     0
1  0         0 47064056 15160924 8751692     0     0    466432  467400   10979  33871   0     7   90   2     0
... many samples skipped ...
r  b      swpd      free     buff     cache    si    so      bi     bo      in     cs    us    sy    id    wa    st
1  0         0  1940860 37169660 30761252     0     0    455680  459200   14158  27724   0     6   91   2     0
0  1         0   920864 37665532 31257060     0     0    495744  500200   14708  28307   0     6   92   2     0
2  0         0  3304476 35016696 31506636     0     0    248832  250136   10781  14505   0     9   89   1     0
1  0         0  9349772 28971956 31506812     0     0         0      0     3503    294   0     6   94   0     0
4  0         0 14827068 23494436 31506704     0     0         0      0    11120    181   0    14   86   0     0
```

# disk stats?

```
# pt-diskstats
```

```
#ts device wr_s wr_avkb wr_mb_s wr_mrg wr_cnc wr_rt busy in_prg
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 595.0 455.6 264.7 0% 1.3 2.2 20% 0
{1} sda5 950.5 456.9 424.1 0% 1.7 1.8 29% 0
{1} sda5 1031.0 456.0 459.1 0% 1.9 1.8 32% 0
{1} sda5 1027.4 456.0 457.5 0% 1.9 1.8 32% 0
{1} sda5 1027.4 456.0 457.5 0% 1.9 1.8 32% 0
... many samples skipped ...
{1} sda5 1045.2 456.0 465.4 0% 1.7 1.7 32% 0
{1} sda5 1072.6 456.4 478.0 0% 2.0 1.9 33% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
{1} sda5 0.0 0.0 0.0 0% 0.0 0.0 0% 0
```

# Let's call sync

```
# sync &
```

```
... and while sync is running:
```

```
# vmstat 1
```

```
procs      -----memory-----  ---swap---  -----io-----  --system--  ---cpu----
 r  b    swpd    free    buff    cache    si  so    bi    bo    in  cs  us  sy  id  wa  st
 0  0      0 31864892 6457368 31506576    0  0     0     0   626 193  0  0 100  0  0
 1  0      0 31865272 6457372 31506572    0  0    28 139452 1521 305  0  2  97  0  0
 0  1      0 31863884 6457372 31506608    0  0     0 855552 4480 264  0  2  93  5  0
... 8 samples skipped ...
 0  1      0 31864424 6457372 31506632    0  0     0 656384 3135 237  0  0  92  7  0
 0  0      0 31864556 6457372 31506636    0  0     0 174909 3556 260  0  2  97  1  0
 0  0      0 31864572 6457372 31506636    0  0     0     0   629 162  0  0 100  0  0
 1  0      0 31864944 6457372 31506640    0  0     0     0   566 149  0  0 100  0  0
 0  0      0 31866232 6457372 31506640    0  0     0     0   464 140  0  0 100  0  0
```

# Small file anyone?

```
# dd if=/dev/zero of=/data/opt/testfile.bin bs=4096 count=262144
1073741824 bytes (1.1 GB) copied, 0.957317 s, 1.1 GB/s
```

```
# vmstat 1
```

```
procs -----memory----- ---swap-- ----io--- --system- -----cpu-----
 r  b   swpd   free   buff   cache   si   so    bi   bo    in   cs  us  sy  id  wa  st
 0  0     0 64976544 6458268  49352    0    0     0    0   701  178  0  0 100  0  0
 0  0     0 64976584 6458268  49356    0    0     0    0   488  160  0  0 100  0  0
 1  0     0 64491596 6458268  534368    0    0     0    0  4022  191  0  4  96  0  0
 0  0     0 63926396 6458276 1097560    0    0     0   12  3315  160  0  5  94  0  0
```

```
# pt-diskstats
```

```
#ts device      wr_s wr_avkb wr_mb_s wr_mrg wr_cnc      wr_rt busy  in_prg
{1} sda5         0.0   0.0   0.0   0%   0.0       0.0  0%    0
{1} sda5         0.0   0.0   0.0   0%   0.0       0.0  0%    0
{1} sda5         0.0   0.0   0.0   0%   0.0       0.0  0%    0
{1} sda5         0.0   0.0   0.0   0%   0.0       0.0  0%    0
{1} sda5         0.0   0.0   0.0   0%   0.0       0.0  0%    0
```



# Synchronize!

```
# sync &
```

```
...
```

```
# vmstat 1
```

```
procs -----memory----- ---swap-- ----io--- --system- ----cpu-----
 r  b   swpd   free   buff   cache   si   so    bi    bo    in   cs  us  sy  id  wa  st
 0  0     0 63928704 6458288 1097956   0   0     0     0   480  154  0  0 99  0  0
 1  0     0 63928076 6458288 1097960   0   0    0 106628 1765  307  0  1 99  0  0
 0  0     0 63927488 6458288 1097964   0   0    0 942027 4311  273  0  1 93  6  0
 0  0     0 63927480 6458288 1097964   0   0    0    34   819  231  0  1 99  0  0
```

```
# pt-diskstats
```

```
#ts device      wr_s wr_avkb wr_mb_s wr_mrg wr_cnc      wr_rt busy  in_prg
{1} sda5         0.0   0.0   0.0   0%   0.0       0.0 0%    0
{1} sda5        200.1 462.5  90.4  0%   5.7       28.7 7%    0
{1} sda5       1656.0 510.5 825.6  0%  114.2      69.0 85%   0
{1} sda5        166.2 484.9  78.7  2%  16.8      100.9 12%   0
{1} sda5         0.0   0.0   0.0  0%   0.0       0.0 0%    0
```

**Reads are  
CACHED**

# %util explained

```
# pt-diskstats
```

rd_s	rd_avkb	rd_mb_s	rd_mrg	rd_cnc	rd_rt	busy	in_prg
238.4	16.0	3.7	0%	1.0	4.2	100%	1
231.5	16.0	3.6	0%	1.0	4.3	99%	1
246.0	16.0	3.8	0%	1.0	4.1	99%	1

# %util explained

```
# pt-diskstats
```

rd_s	rd_avkb	rd_mb_s	rd_mrg	rd_cnc	rd_rt	busy	in_prg
238.4	16.0	3.7	0%	1.0	4.2	100%	1
231.5	16.0	3.6	0%	1.0	4.3	99%	1
246.0	16.0	3.8	0%	1.0	4.1	99%	1
479.6	16.0	7.5	0%	2.0	4.1	100%	2
472.0	16.0	7.4	0%	2.0	4.2	100%	2
456.2	16.0	7.1	0%	2.0	4.4	100%	2

# %util explained

# pt-diskstats

rd_s	rd_avkb	rd_mb_s	rd_mrg	rd_cnc	rd_rt	busy	in_prg
238.4	16.0	3.7	0%	1.0	4.2	100%	1
231.5	16.0	3.6	0%	1.0	4.3	99%	1
246.0	16.0	3.8	0%	1.0	4.1	99%	1
479.6	16.0	7.5	0%	2.0	4.1	100%	2
472.0	16.0	7.4	0%	2.0	4.2	100%	2
456.2	16.0	7.1	0%	2.0	4.4	100%	2
840.7	16.0	13.1	0%	4.0	4.7	100%	4
857.9	16.0	13.4	0%	4.0	4.6	100%	4
837.9	16.0	13.1	0%	4.0	4.8	100%	4
1375.4	16.0	21.5	0%	8.0	5.8	100%	8
1386.2	16.0	21.7	0%	8.0	5.8	100%	8
1435.3	16.0	22.4	0%	8.0	5.6	100%	8
2147.4	16.0	33.6	0%	16.0	7.4	100%	16
2177.7	16.0	34.0	0%	16.0	7.3	100%	16
2105.1	16.0	32.9	0%	16.0	7.6	100%	16
4057.4	16.0	63.4	0%	63.8	15.7	100%	64
4038.3	16.0	63.1	0%	63.7	15.8	100%	64
4020.8	16.0	62.8	0%	64.3	16.0	100%	64

# Tools?

# vmstat

```
procs -----memory----- ---swap-- ----io----- --system-- -----cpu-----
 r  b   swpd   free   buff  cache   si   so    bi    bo    in    cs  us  sy  id  wa  st
 2  7    140 157964 59124 394676   0    0 12312  5180 4006   9755 32  5 35 28  0
 6  4    140 159260 57188 394416   0    0 11352    16 3324   9965 36  4 40 20  0
 3  7    140 161832 54280 395472   0    0  9292  3648 3522 10218 67  7 12 14  0
24  4    140 160468 54400 396220   0    0 10356  2608 3547 12196 62  5 20 13  0
14  4    140 157368 54604 399152   0    0 11192  2048 2941 11666 58  7 21 14  0
 7  3    140 156360 54656 400092   0    0  7608  2068 4555 10273 43  5 37 15  0
 6  5    140 158828 54680 399180   0    0  8340  3832 3914   9622 54  5 29 13  0
 3  4    140 157996 55084 398408   0    0  8856  8688 3327   9929 43  5 35 16  0
 1  5    140 173740 55168 383104   0    0  7488  1992 2338 10352 46  6 34 14  0
 1  3    140 170748 55464 384260   0    0  6840  2560 2545 12245 57  6 25 13  0
 2  8    140 168516 55692 387740   0    0  4952  3112 2085 11066 42  5 41 12  0
 2  5    140 167236 55840 388140   0    0  6352    716 2134 11347 42  5 35 18  0
 4  6    140 165056 55984 389764   0    0  5896  2036 2043   8925 35  3 38 24  0
 7  6    140 164024 56032 390944   0    0  5428  1672 1944   9485 43  4 34 20  0
 6  4    140 161500 56100 393116   0    0  4496  1588 2038 13300 54  5 26 15  0
 3  6    140 158312 56184 395904   0    0  4824  2204 1989 12303 59  6 15 20  0
 4  6    140 158068 56300 396412   0    0  5632  4208 2238 10127 50  4 24 22  0
 5  8    140 155828 56460 398960   0    0  4104  2152 1815   7972 28  2 43 26  0
 4  4    140 158680 55092 395920   0    0  7616  1764 2010   8047 27  2 54 17  0
10  5    140 157588 55156 396780   0    0  3076  2208 1906 10372 36  3 31 29  0
 5  4    140 158332 55200 396576   0    0  5008  1868 1936   9832 46  3 35 16  0
```

**IT'S JUST**  
**a system.**



?



**aurimas, percona.com**

**We're Hiring! [www.percona.com/about-us/careers/](http://www.percona.com/about-us/careers/)**