Stainless Steel Elephant

Standing on the shoulders of giants

Alex Chistyakov, Principal Engineer at Git in Sky
Feb 05 2016, PgConf.Russia 2016

Git in Sky
TUNING YOUR INFRASTRUCTURE

Who we are

- A small consulting company from SPb., Russia
- Web ops engineers, performance engineers
- Automation engineers
- PostgreSQL fans

Git in Sky
TUNING YOUR INFRASTRUCTURE

Who are you?

- DBAs?
- DBDs?
- DB(.*)s?
- PostgreSQL fans?
- Performance engineering, anyone?

Git in Sky
TUNING YOUR INFRASTRUCTURE

OK, what is going on here?

- FreeBSD!

Git in Sky
TUNING YOUR INFRASTRUCTURE

OK, what is going on here?

- FreeBSD!
- Not scared? Let's try again!

Git in Sky
TUNING YOUR INFRASTRUCTURE

OK, what is going on here?

- FreeBSD!
- Not scared? Let's try again!
- ZFS!

Git in Sky
TUNING YOUR INFRASTRUCTURE

OK, what is going on here?

- FreeBSD!
- Not scared? Let's try again!
- ZFS!
- Hmm, you still there?

Git in Sky
TUNING YOUR INFRASTRUCTURE

OK, what is going on here?

- FreeBSD!
- Not scared? Let's try again!
- ZFS!
- Hmm, you still there?
- DragonFly BSD! HAMMER!

Git in Sky
TUNING YOUR INFRASTRUCTURE

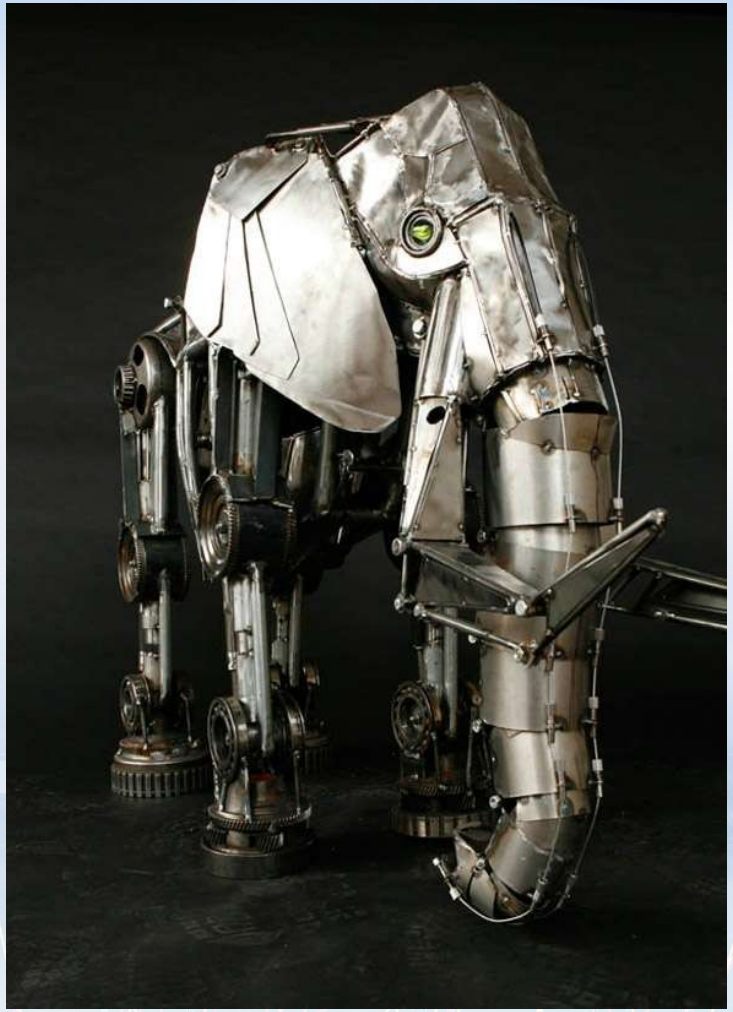
OK, what is going on here?

- FreeBSD!
- Not scared? Let's try again!
- ZFS!
- Hmm, you still there?
- DragonFly BSD! HAMMER!
- Okay, enough for you

Git in Sky
TUNING YOUR INFRASTRUCTURE

Hardware configuration

- Dell R430
- 32Gb RAM
- PERC H730 mini
- Two Edge Boost Pro P SSDs in RAID0
- 2 x Xeon(R) CPU E5-2630 v3 @ 2.40GHz



Git in Sky
TUNING YOUR INFRASTRUCTURE

Ready, set, go!

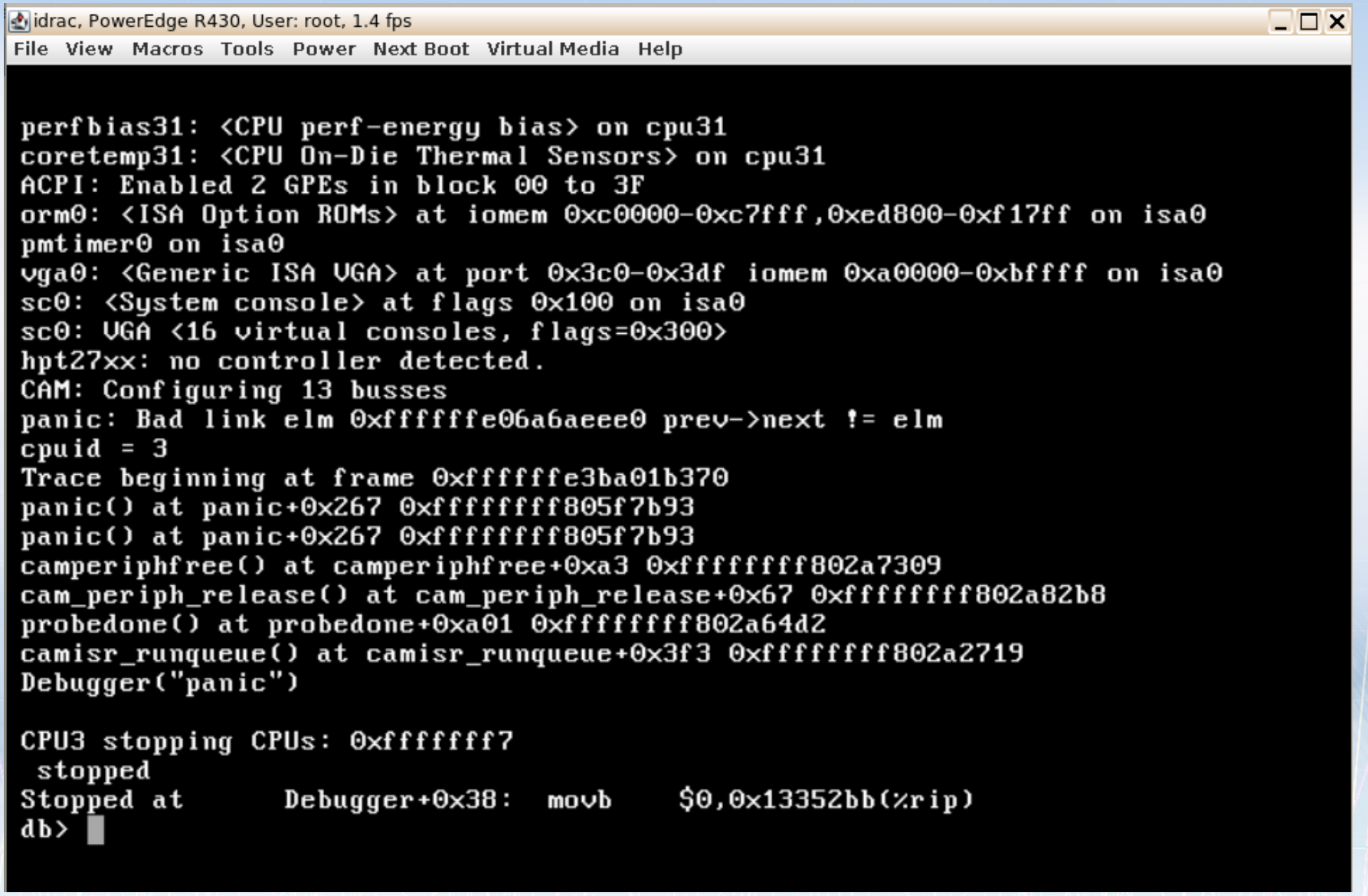
- BTW, the plan is:
- Install something very elite
- Get great results
- Install something less elite
- Get not so great results
- Compare and swap
- PROFIT!

Git in Sky
TUNING YOUR INFRASTRUCTURE

DragonFly BSD 4.4.1 (the latest)

- It's dead, Jim!
- Was not able

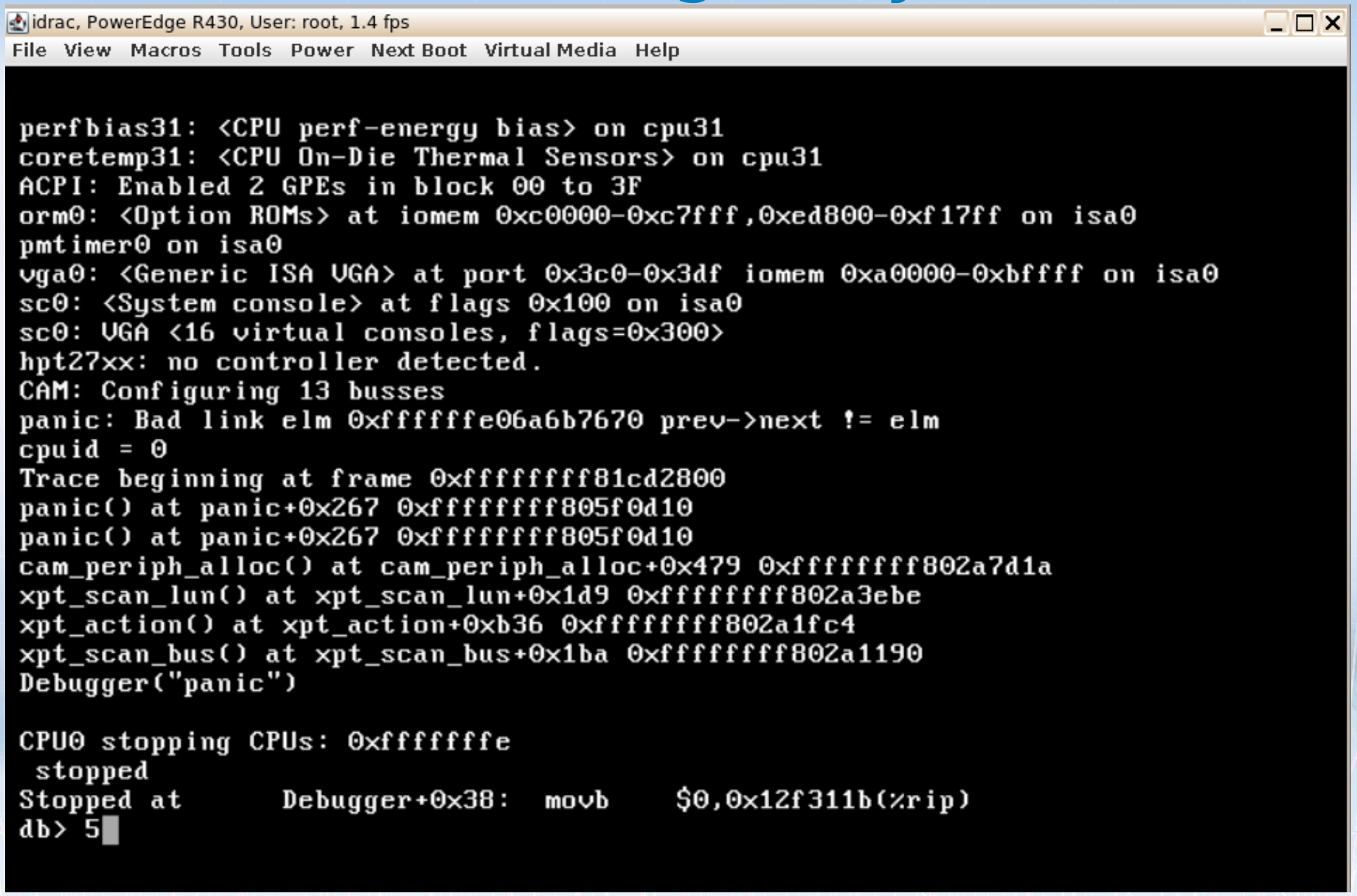
  to even install it



Git in Sky
TUNING YOUR INFRASTRUCTURE

DragonFly BSD 4.2.4

- DOA too



Git in Sky
TUNING YOUR INFRASTRUCTURE

DragonFly BSD 4.0.6

- Hoorah!
- Well, that's elite enough
- [root@dfbsd /usr]# uname -a

  DragonFly dfbsd.gitinsky.com 4.0-RELEASE DragonFly v4.0.6-RELEASE #0: Fri Jun 12 19:57:23 EDT 2015 root@www.shiningsilence.com:/usr/obj/home/justin/release/4_0/sys/X86_64_GENERIC x86_64

  [root@dfbsd /usr]#

Git in Sky
TUNING YOUR INFRASTRUCTURE

Now you have two problems

- Okay, what's next?
- “pgbench -i -s 1000 --foreign-keys pgbench” to load some data
- “pgbench -T 300 -P 10 -c N -j N -r pgbench” to run some tests
- A carefully trained monkey to interpret the results

Git in Sky
TUNING YOUR INFRASTRUCTURE

Loading.....

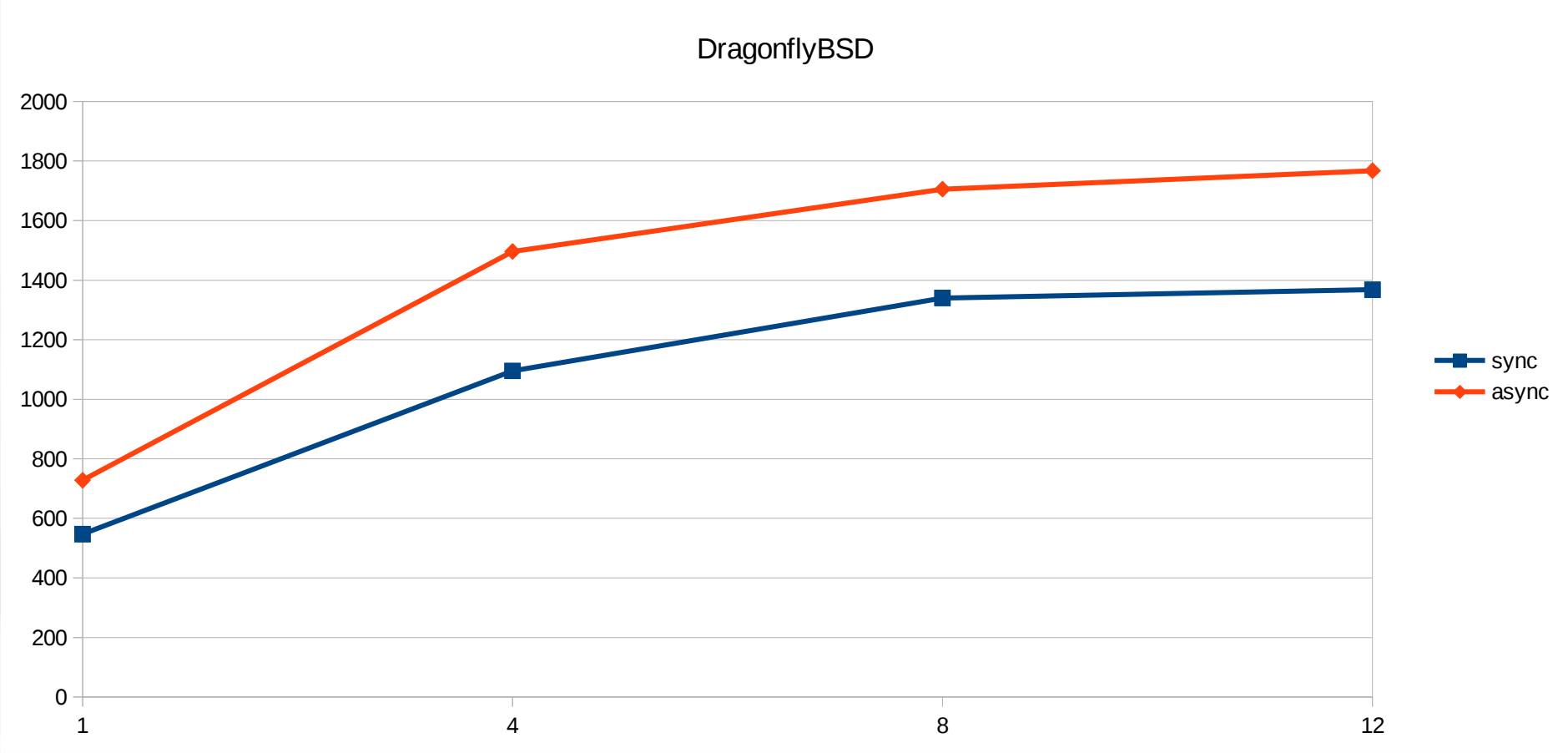
- date; pgbench -i -s 1000 --foreign-keys pgbench; date

  Fri Feb 5 03:33:28 MSK 2016
  creating tables...
  100000000 of 100000000 tuples (100%) done (elapsed 257.56 s, remaining 0.00 s)
  vacuum...
  set primary keys...
  set foreign keys...
  done.
  Fri Feb 5 03:43:30 MSK 2016

Git in Sky
TUNING YOUR INFRASTRUCTURE

Here we go!

- X – num of clients/threads, Y – transactions per second



Git in Sky
TUNING YOUR INFRASTRUCTURE

A note on observability tools

- I'm in your gdb killing your backendz!
- [root@dfbsd /home/chistyakov]# gdb -ex "set pagination 0" -ex "thread apply all bt" -batch -p 436012

  Couldn't get registers: Device busy.

  Couldn't get registers: Device busy.

  Couldn't get registers: Device busy.

  Thread 1 (process 436012):

  Couldn't get registers: Device busy.

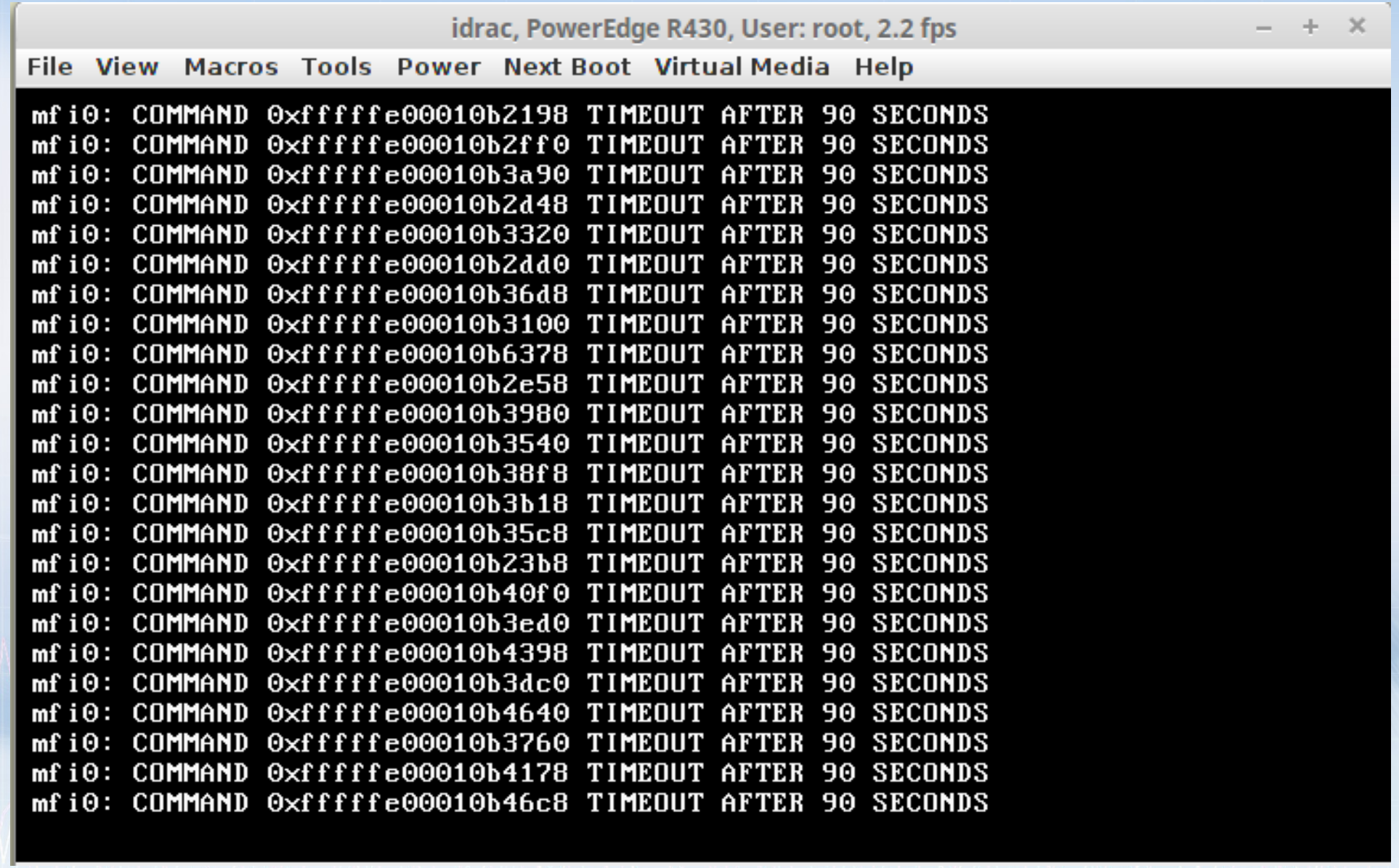
  Quitting: ptrace: Device busy.

  [root@dfbsd /home/chistyakov]#

Git in Sky
TUNING YOUR INFRASTRUCTURE

FreeBSD 10.2

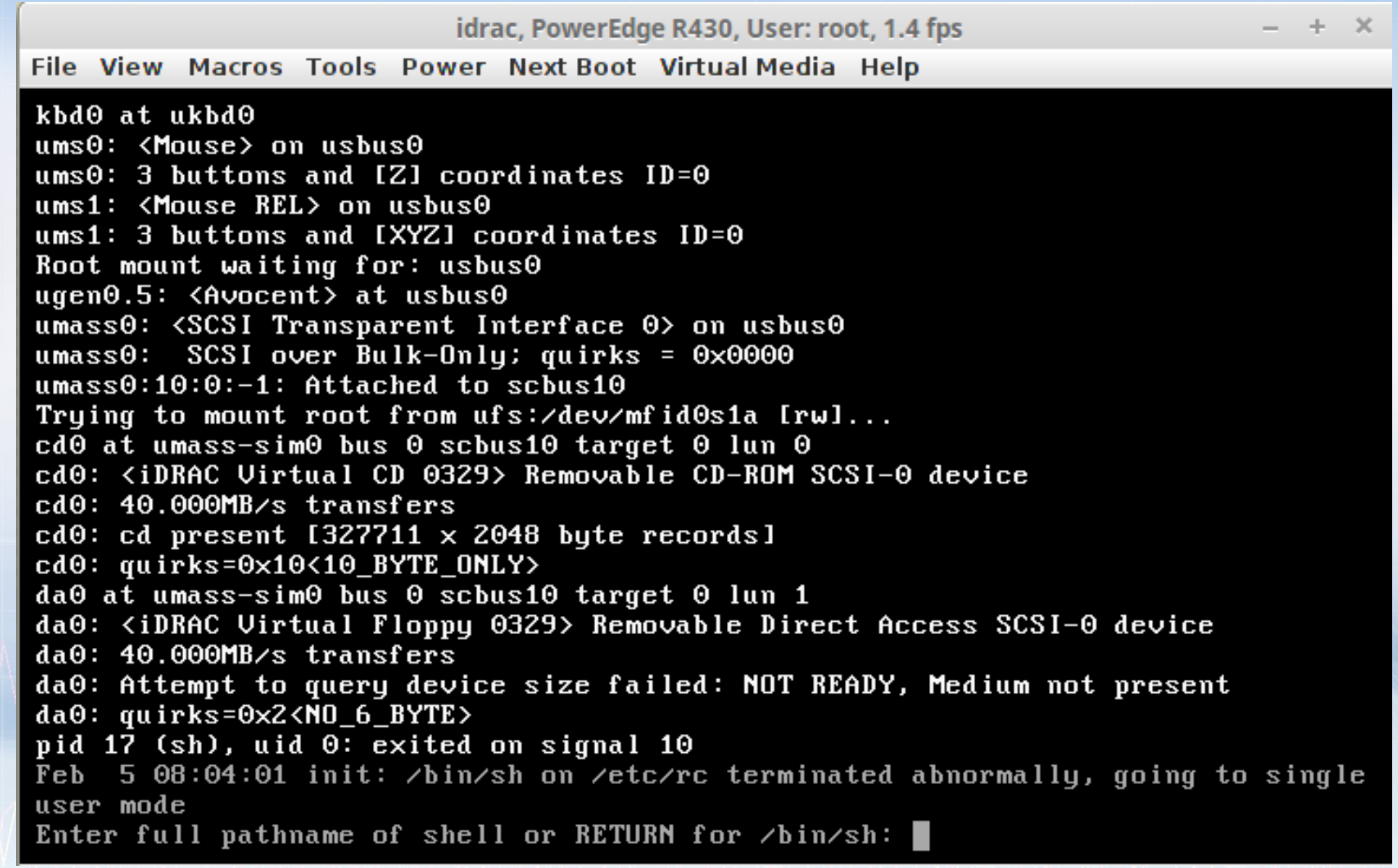
- BSD guys, what is wrong with you?



Git in Sky
TUNING YOUR INFRASTRUCTURE

FreeBSD 9.3

- This is how a deferred success looks like



Git in Sky
TUNING YOUR INFRASTRUCTURE

Okay, SmartOS then

- 100000000 of 100000000 tuples (100%) done (elapsed 107.03 s, remaining 0.00 s).
  vacuum...
  set primary keys...
  set foreign keys...
  done.

  real   4m27.237s
  user   0m23.381s
  sys   0m2.118s
  bash-4.1$

Git in Sky
TUNING YOUR INFRASTRUCTURE

A note on observability tools

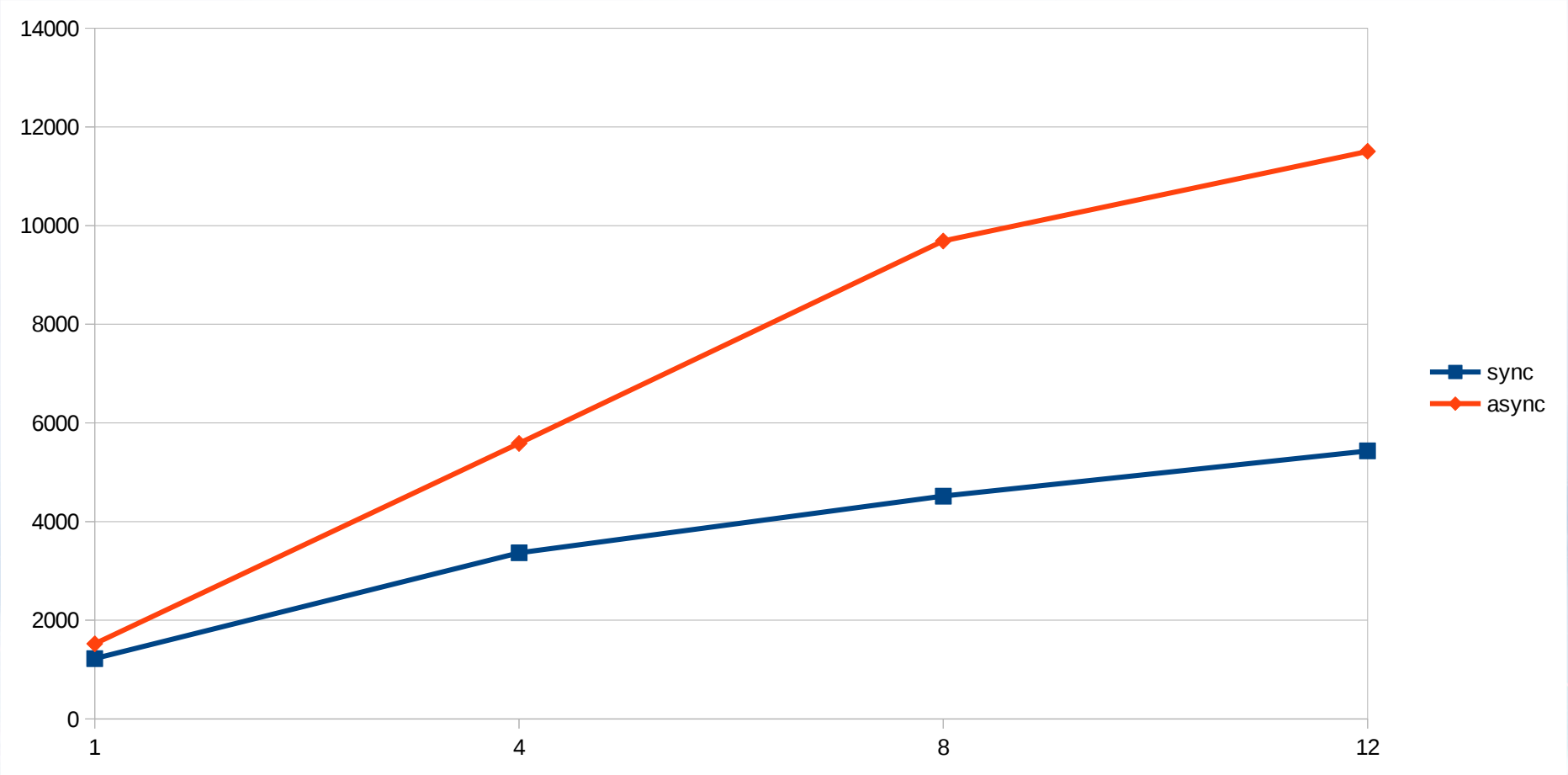
- zpool iostat 1

```
zones       25,3G   419G      0  4,83K      0   417M
zones       25,3G   419G      0  4,63K      0   400M
zones       25,3G   419G      0  4,58K      0   386M
zones       25,4G   419G      0  4,55K      0   366M
zones       25,4G   419G      0  4,57K      0   364M
zones       25,4G   419G      0  4,51K      0   378M
zones       25,4G   419G      0  4,20K      0   358M
zones       25,4G   419G      0  4,83K      0   422M
zones       25,5G   419G      0  4,32K      0   360M
```

Git in Sky
TUNING YOUR INFRASTRUCTURE

Sync/async, SmartOS*

- * This chart is a lie



Git in Sky
TUNING YOUR INFRASTRUCTURE

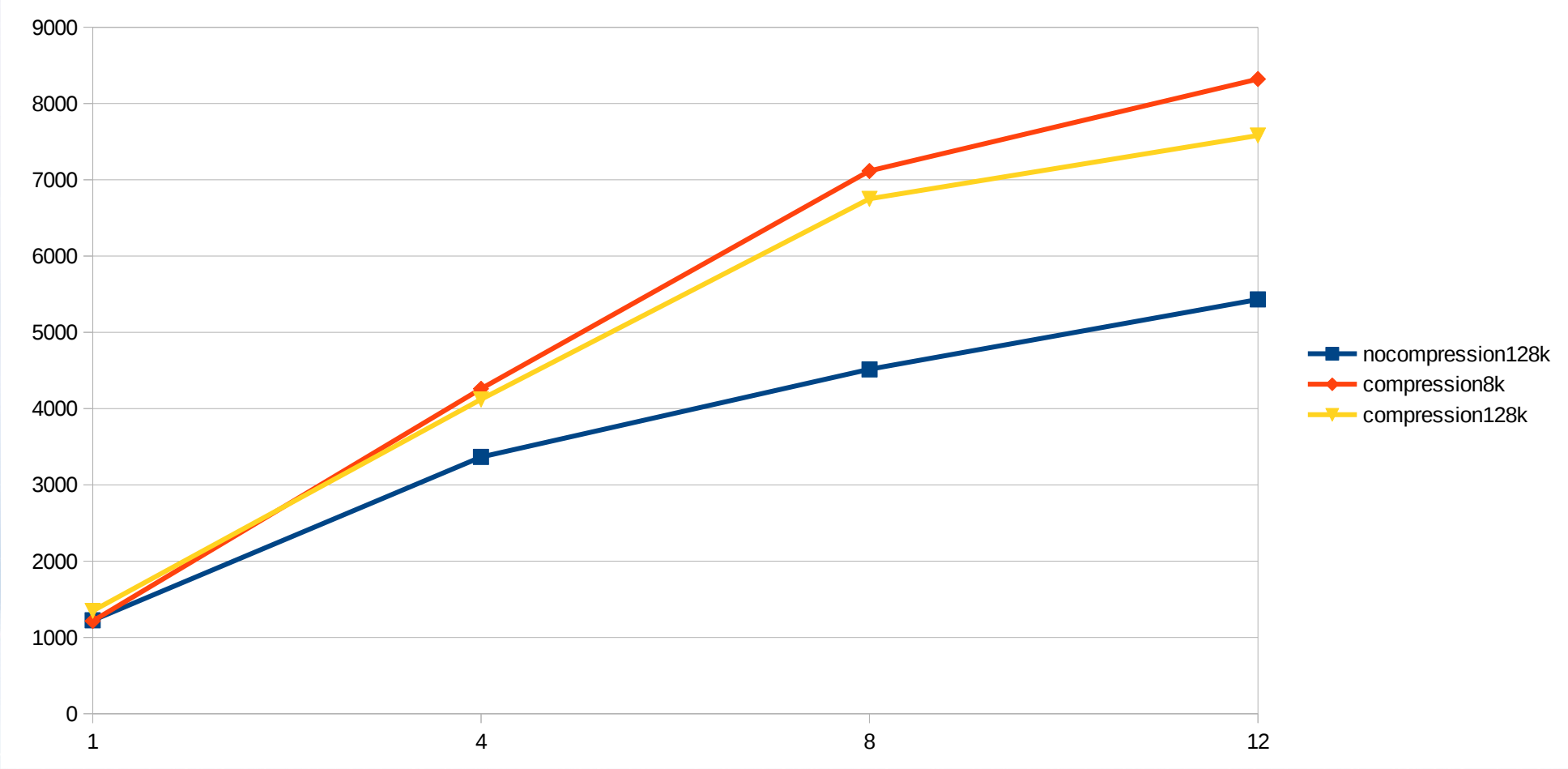
Lies, damned lies and LZO

- Compression MUST be turned on, period
- ZFS record size: 128K by default
- PostgreSQL block: 8K
- Compressing 128K blocks is more efficient
- Random reads: 8K → 128K amplification
- Fight!

Git in Sky
TUNING YOUR INFRASTRUCTURE

The joy of ZFS

- * This chart is probably a lie



Git in Sky
TUNING YOUR INFRASTRUCTURE

Guess which one is 128K?

- zones/var refcompressratio 5.20x -
  zones/var written 3,31G -
  zones/var logicalused 16,8G -
  zones/var logicalreferenced 16,8G -

- zones/var refcompressratio 11.64x -
  zones/var written 1,27G -
  zones/var logicalused 14,7G -
  zones/var logicalreferenced 14,7G -

Git in Sky
TUNING YOUR INFRASTRUCTURE

Ubuntu 14.04 LTS

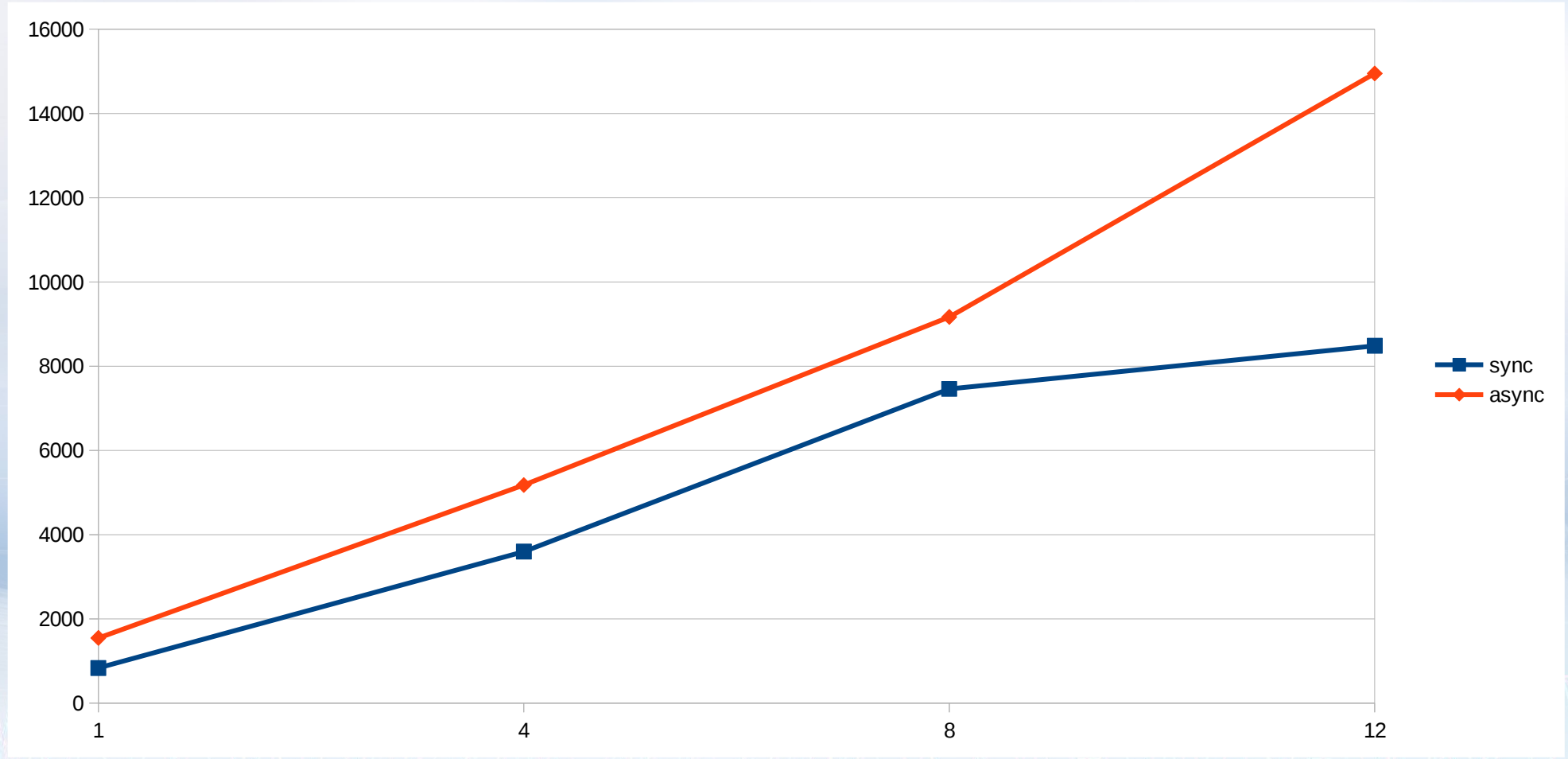
- 100000000 of 100000000 tuples (100%) done (elapsed 65.91 s, remaining 0.00 s)
  vacuum...
  set primary keys...
  set foreign keys...
  done.

  real 2m54.867s
  user 0m27.372s
  sys 0m0.872s
  postgres@ubuntu:~$

Git in Sky
TUNING YOUR INFRASTRUCTURE

Sync/async, Ubuntu/XFS

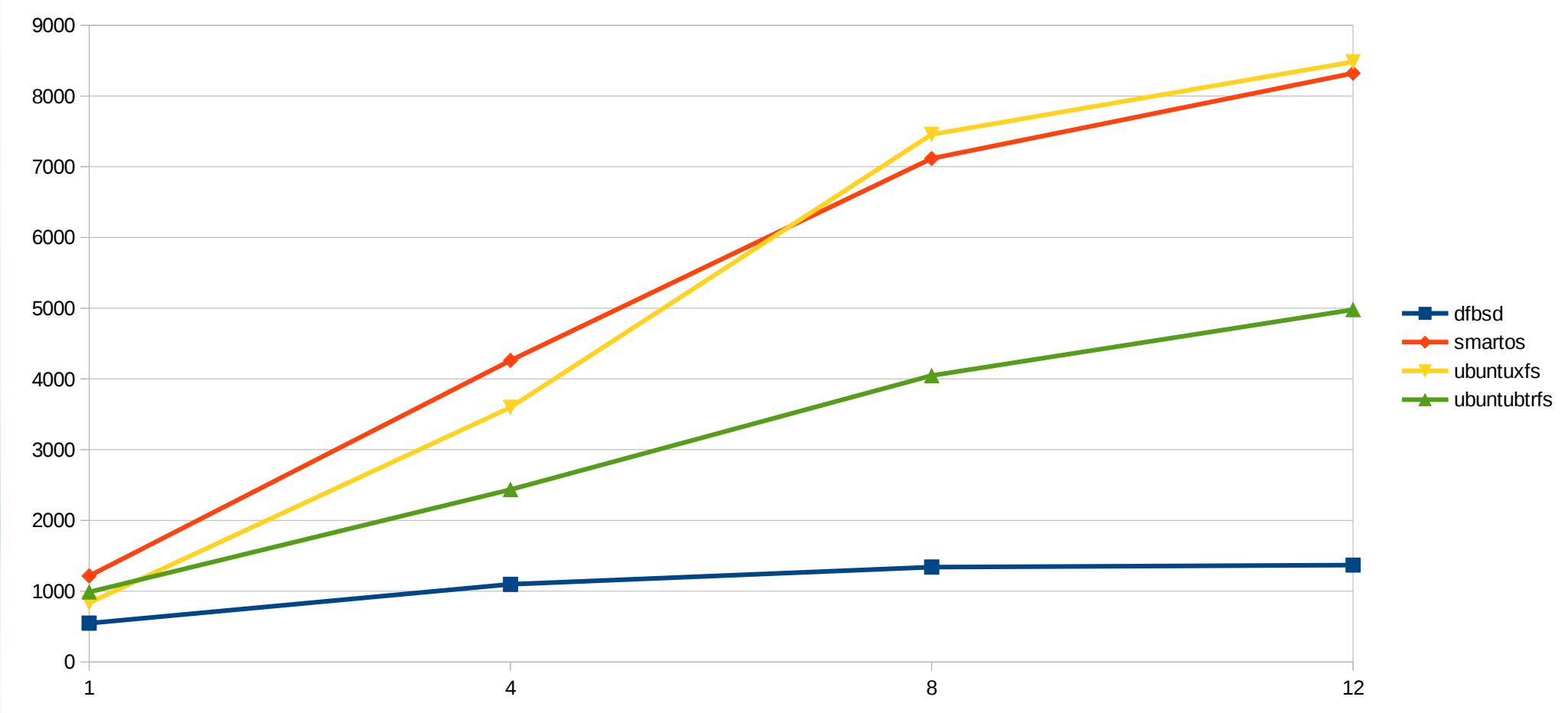
- This chart is not a lie



Git in Sky
TUNING YOUR INFRASTRUCTURE

Synchronous commit

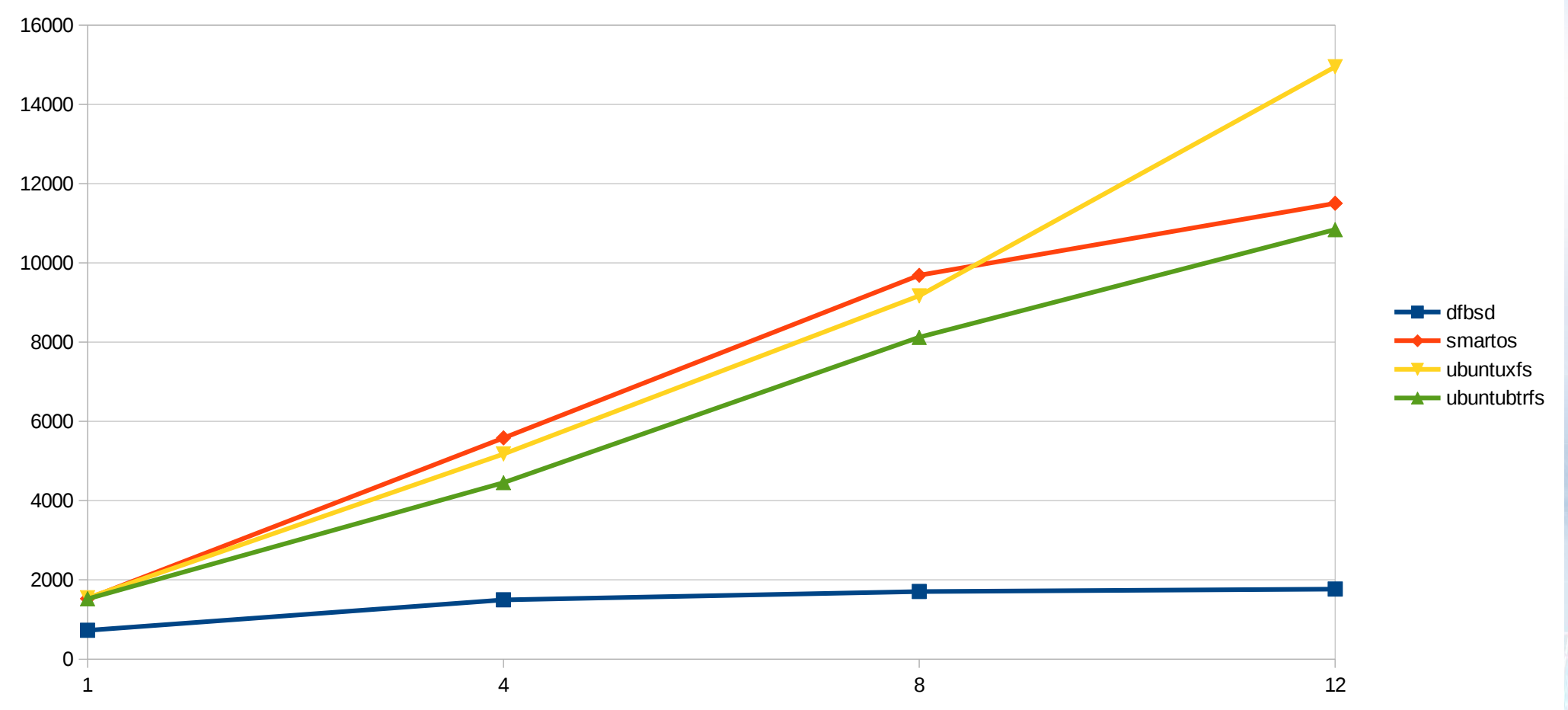
- SmartOS/ZFS (compression, 8K block) – Ubuntu/XFS : 1 - 1



Git in Sky
TUNING YOUR INFRASTRUCTURE

Asynchronous commit

- And, finally, we have a winner! Ubuntu/XFS



Git in Sky
TUNING YOUR INFRASTRUCTURE

Conclusions

- Hammers don't fly at decent speed
- FreeBSD is very elite
- But you still can use ZFS if you are not elite enough for FreeBSD
- Ubuntu and XFS are not even close to be elite but are very fast
- SmartOS? Never heard of it!

Git in Sky
TUNING YOUR INFRASTRUCTURE

Q & A

- - Why did you tweak ZFS but did not tweak BTRFS and HAMMER?
- - Because I'm a ZFS zealot, you heretics!

Git in Sky
TUNING YOUR INFRASTRUCTURE

That's all folks!

- Thank you!
- alex@gitinsky.com
- http://gitinsky.com
- http://meetup.com/DevOps-40