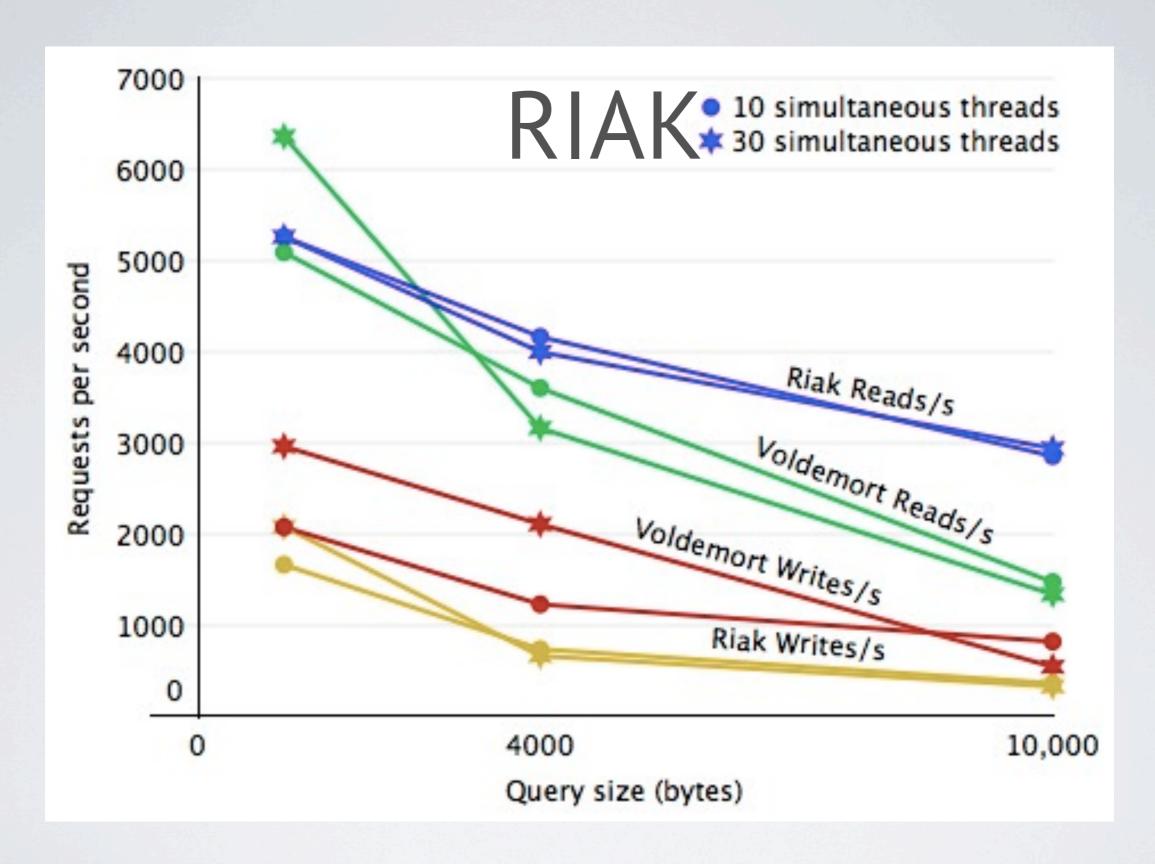# AMAZON AWS + RIAK

Lev Walkin, Denis Titoruk
(Devil's advocates)

• http://lionet.livejournal.com/tag/riak

HISTORY

- March 2010 — 10 Amazon EC2 Medium instances (0.9)

- January 2011 — Upgrade to 10 EC2 Large instances (0.13)

- Pre-launch data: 50 mln entries, 200 GB Data.

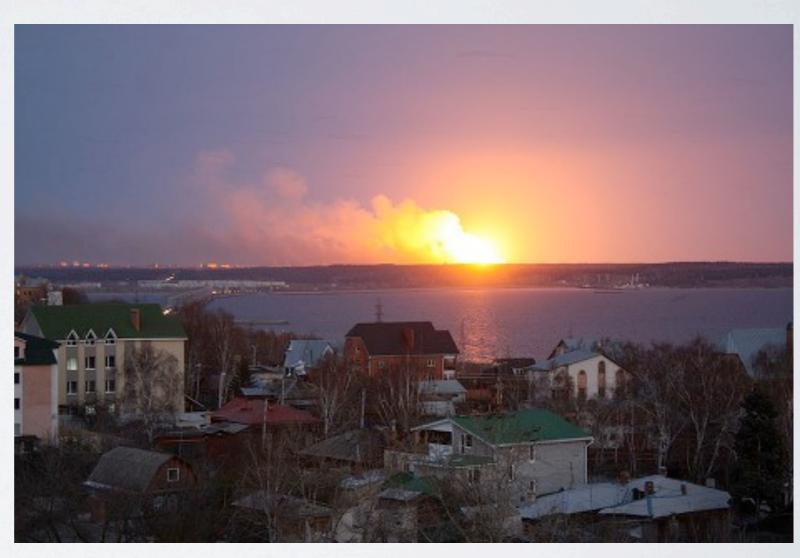- Ring size 512 (up from defalut 64), InnoDB @ EBS

RIAK QUIRKS

- Delete operation is not consistent (ordering race).

- Vnode partitions get transferred in parallel during rebalancing.

- Ring sizes determines scale and performance.

INNODB QUIRKS

- InnoDB max key size 256 bytes.

- InnoDB runs faster with monotonously increasing keys.

- Riak buckets use file descriptors geometrically.

BITCASK QUIRKS

- Maturity problems →

EXPERIENCE

- Riak 0.9 upgrade story.

- Erlang native interface to Riak is lagging behind.

- [Number of buckets *] number of vnodes = write contention.

- Parallel rebalancing greatly affects performance.

- Read+Write serializes.

EC2 + EBS + RIAK

- − Flaky operation on lower-grade boxes (MTBF <1 month)

- + Easy box upgrades instead of cluster rebalancing.

- + EBS "volume cloning" allows upgrade rehearsing.

RECOMMENDATIONS

- Amazon EC2 upgrades are the easiest way to scale.

- Cluster expansions are tricky.

- Keep data in memory (surprise!)

- Plan carefully: (N≠3?, ring size≠64, key structure (≠md5) and key length).

- Test on realistic R/W mix with realistic key distributions and data set size.

WRAP-UP

- Riak works! Plus, great community and momentum!

- Reasonable handling of overload: does not blow up.

- Cluster expansions might b=2 effective multi-hour downtime.

- BitCask is still risky.