

# Cassandra Hackathon

Not only bad news

Daniel Doubleday

[dd@smeet.com](mailto:dd@smeet.com)

Smeet Communications GmbH

# Smeet: Some Numbers

- Flash virtual world  
(Rooms, Avatars, Chat, Games etc)
- 10M Users
- 100k Users in hot data set
- ca. 50 Servers

# Use Cases

- 5 Nodes: Read/Write Main Cluster (100 reads, 30 writes per node)
  - User Data (small) (50GB)
  - Blobs (larger: 1 - 5k) (200GB)
- 2 Nodes: Write Only Log Cluster
  - Chatlogs, OLAP Biz Data
- X Nodes: Flume Sink (In Progress)

# Hardware

- 6 - Core AMD
- 16 GB RAM (max 32)
- 1 x 1 TB System Drive
- 2 x 500 GB Velociraptor Raid 0
- 1 x 2 TB Backup Drive
  
- Want to switch to Intel 600GB SSDs

# Random Issues

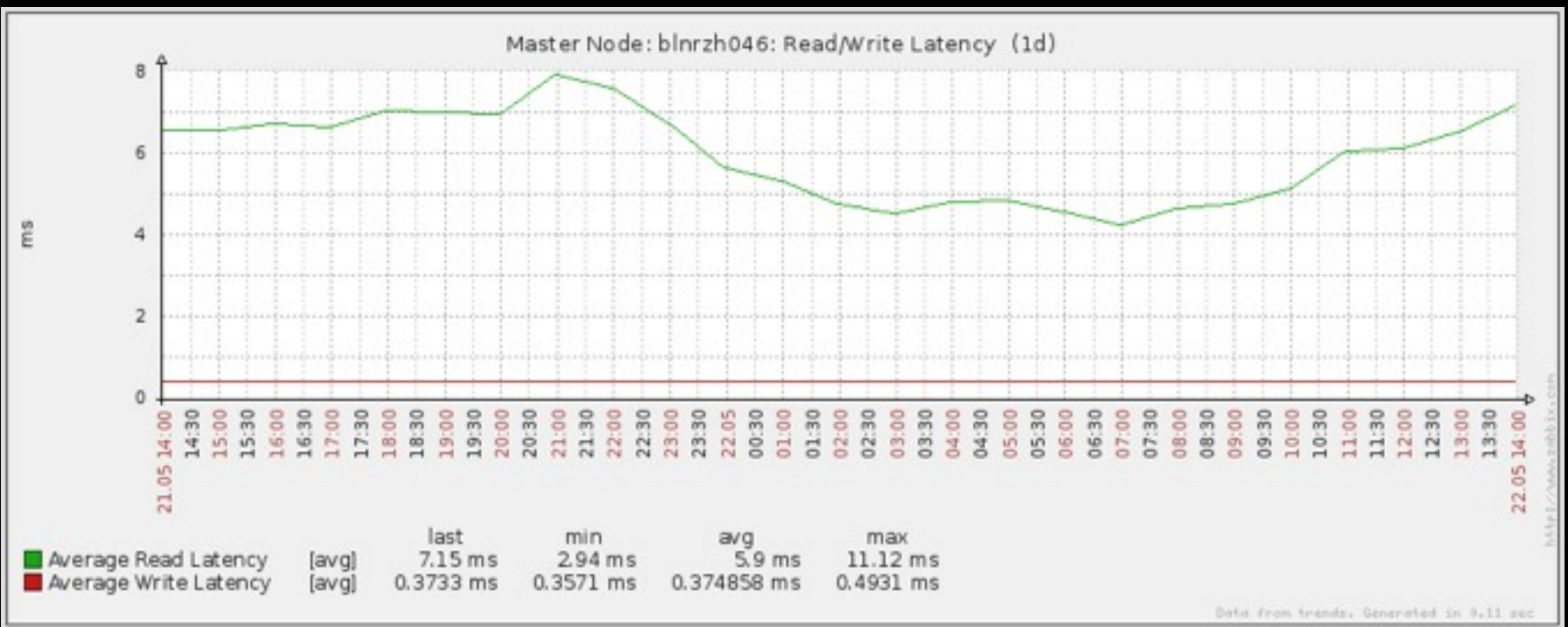
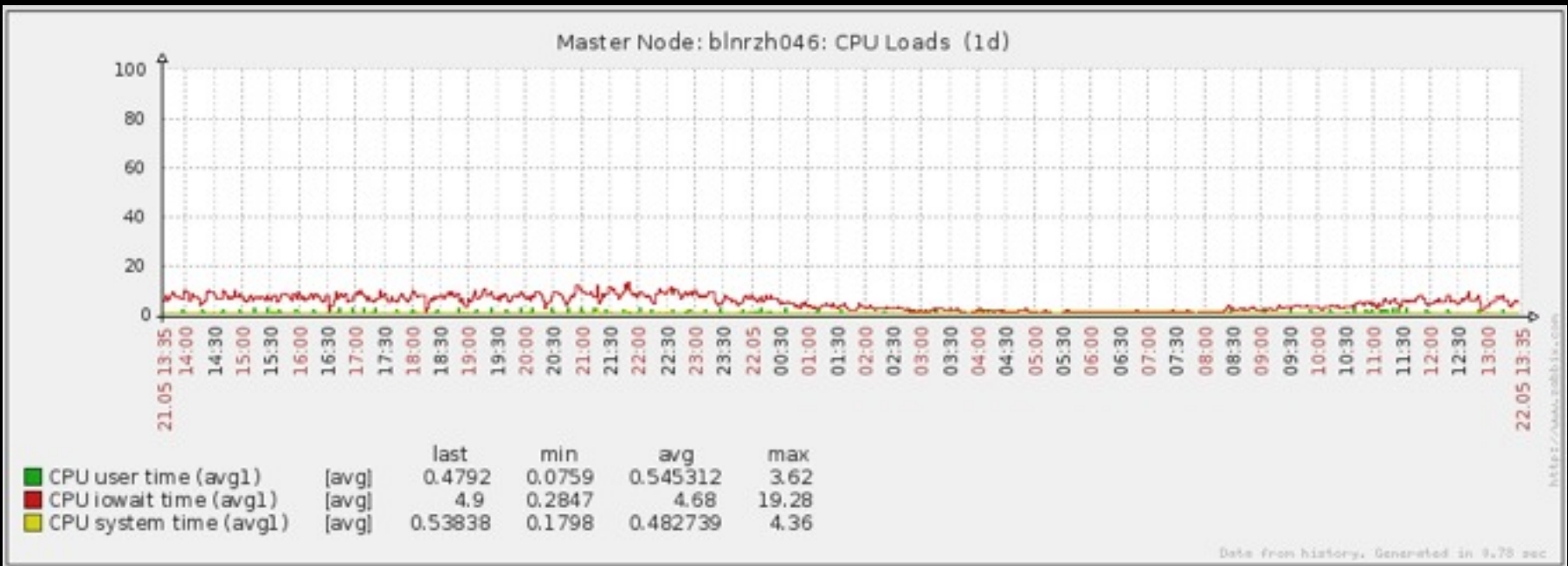
- Quorum Read Load
- Compaction / Repair
- Atomicity: Data Model
- Big Data Updates

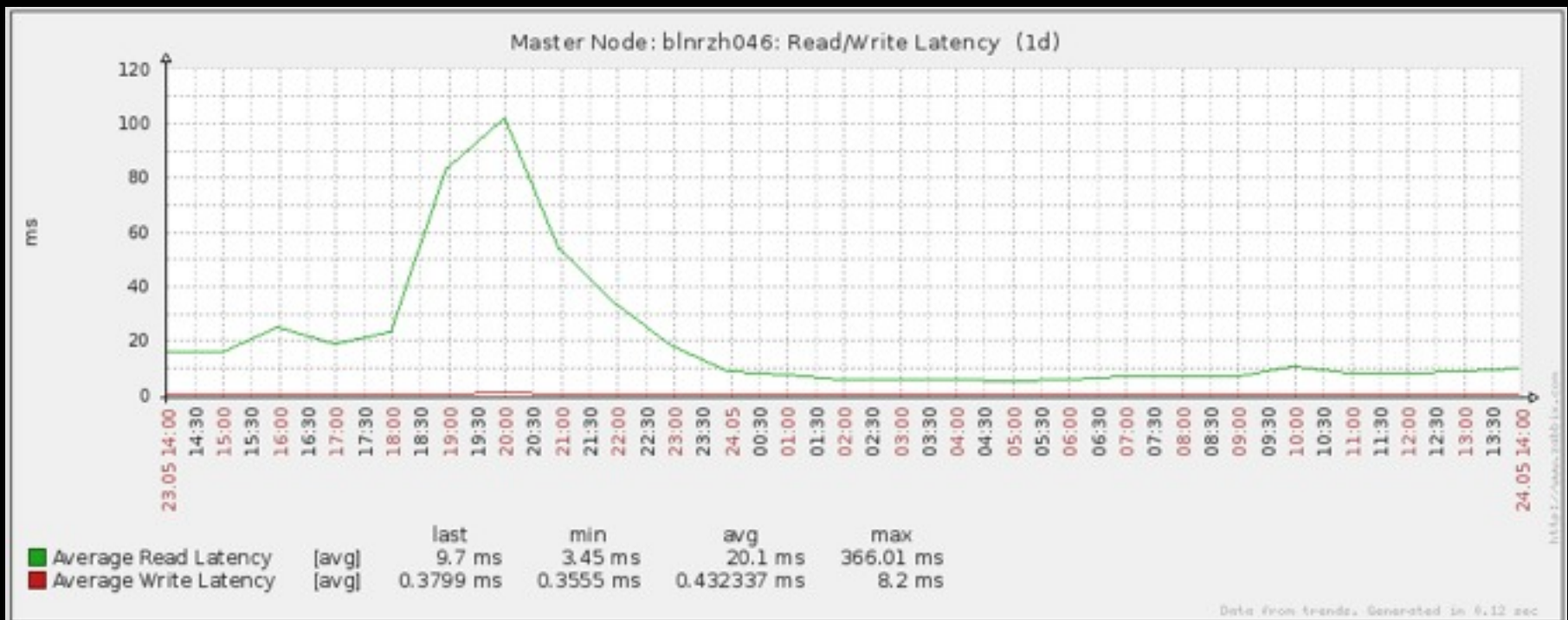
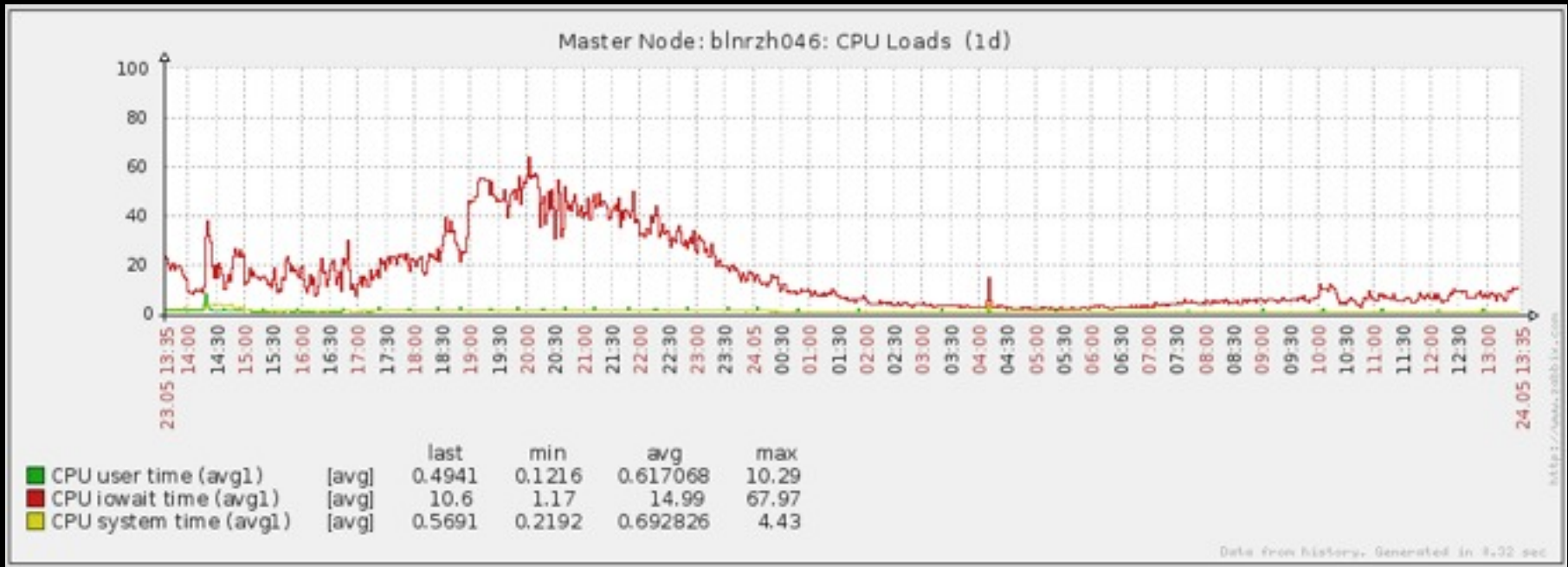
# Quorum Read / Write

- Read Repair can kill an overloaded system
- Use only RF nodes to read from
- ReadRepair Probability: 0

# IO Load

- I'm doing good - right?



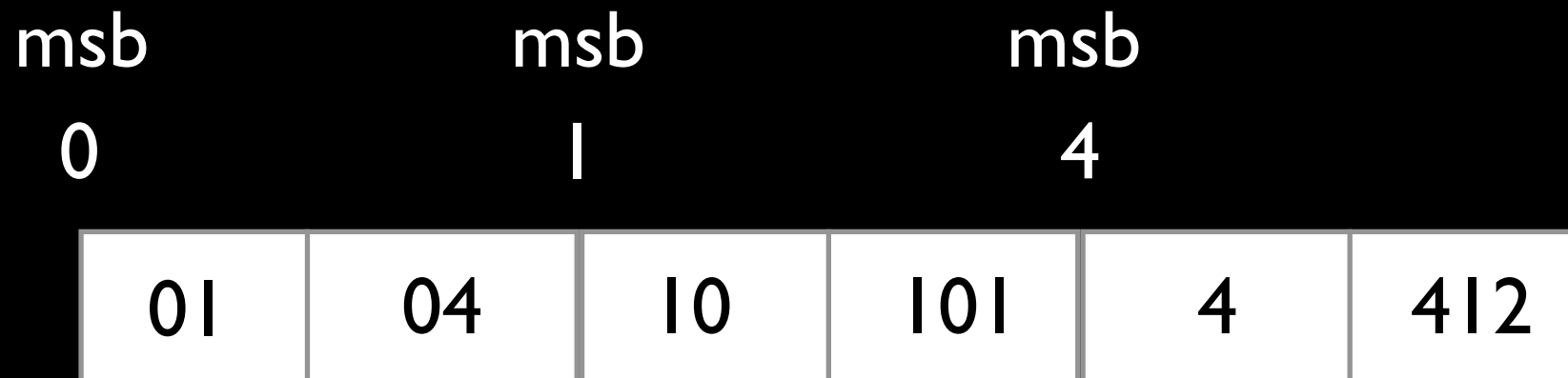


# Atomicity

- Probability of inconsistencies increase when hitting different sections in the ring
  - Use same row key
  - TODO: create custom replication strategy and annotated token to overwrite key token

# Prefix Column Keys

- We always use `ByteType`



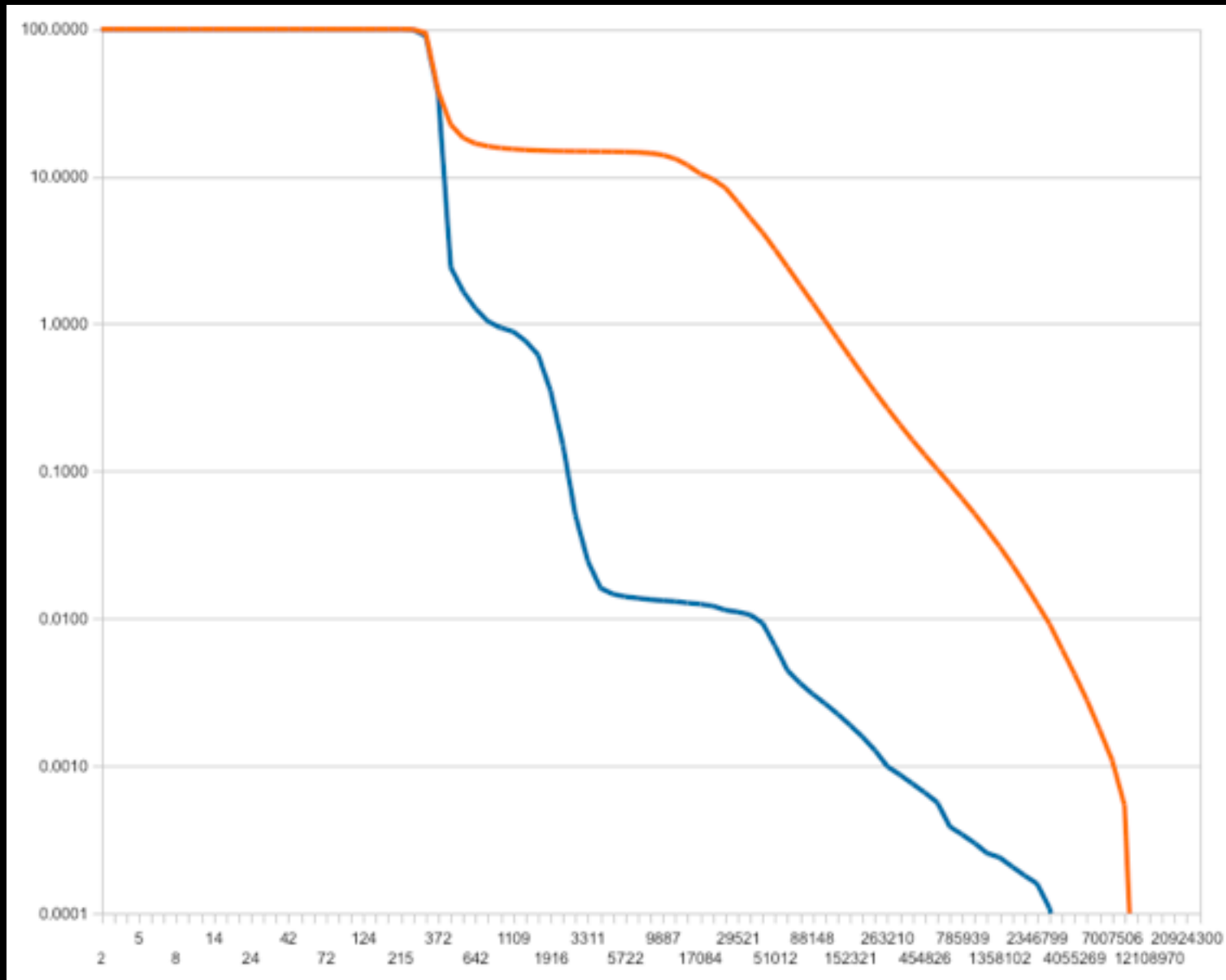
- Slicing / Semantic Grouping

# Big Data Updates

- Many Updates on large CFs are bad
- Our solution: create WORM CF

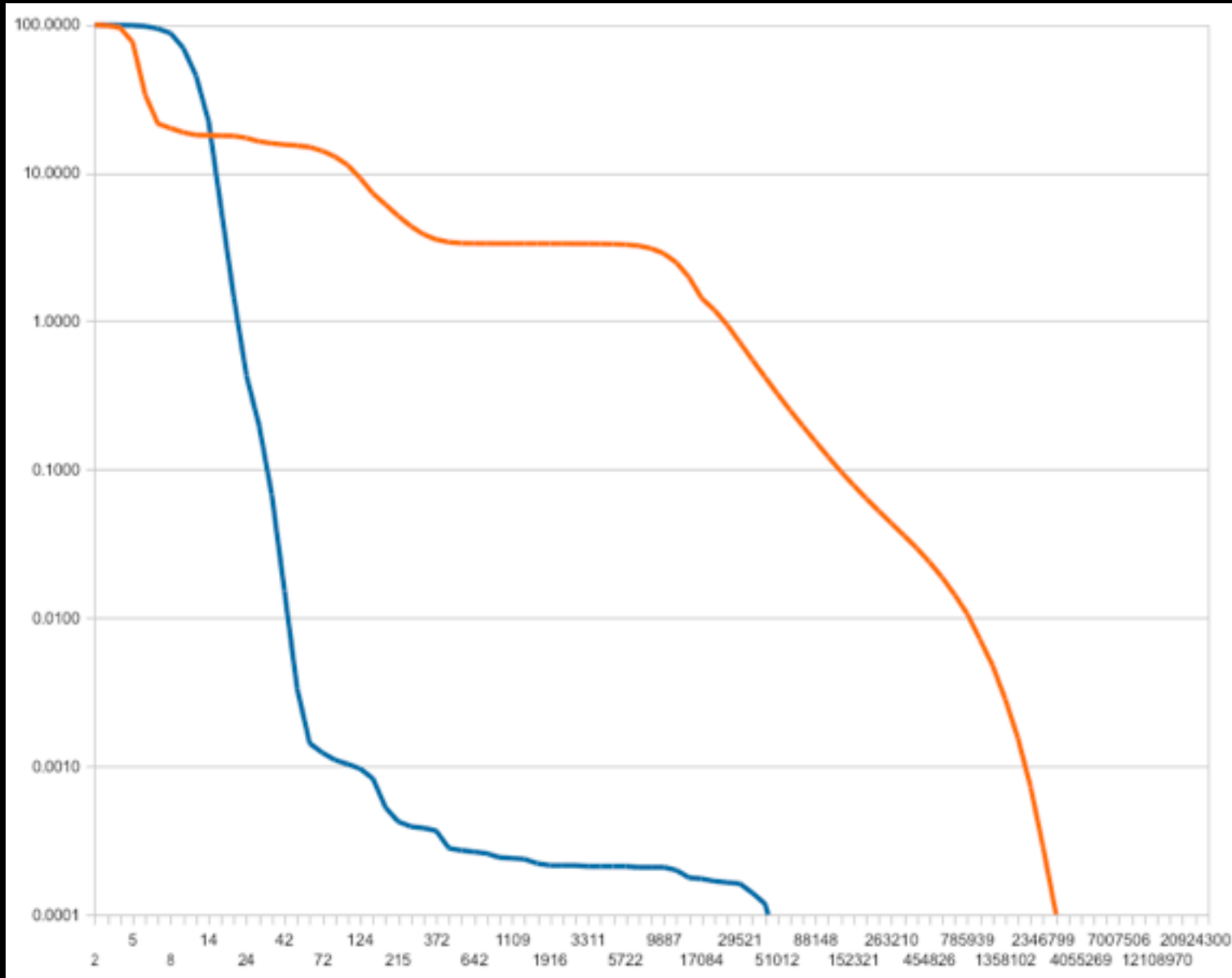
# Latencies

## Storage Proxy



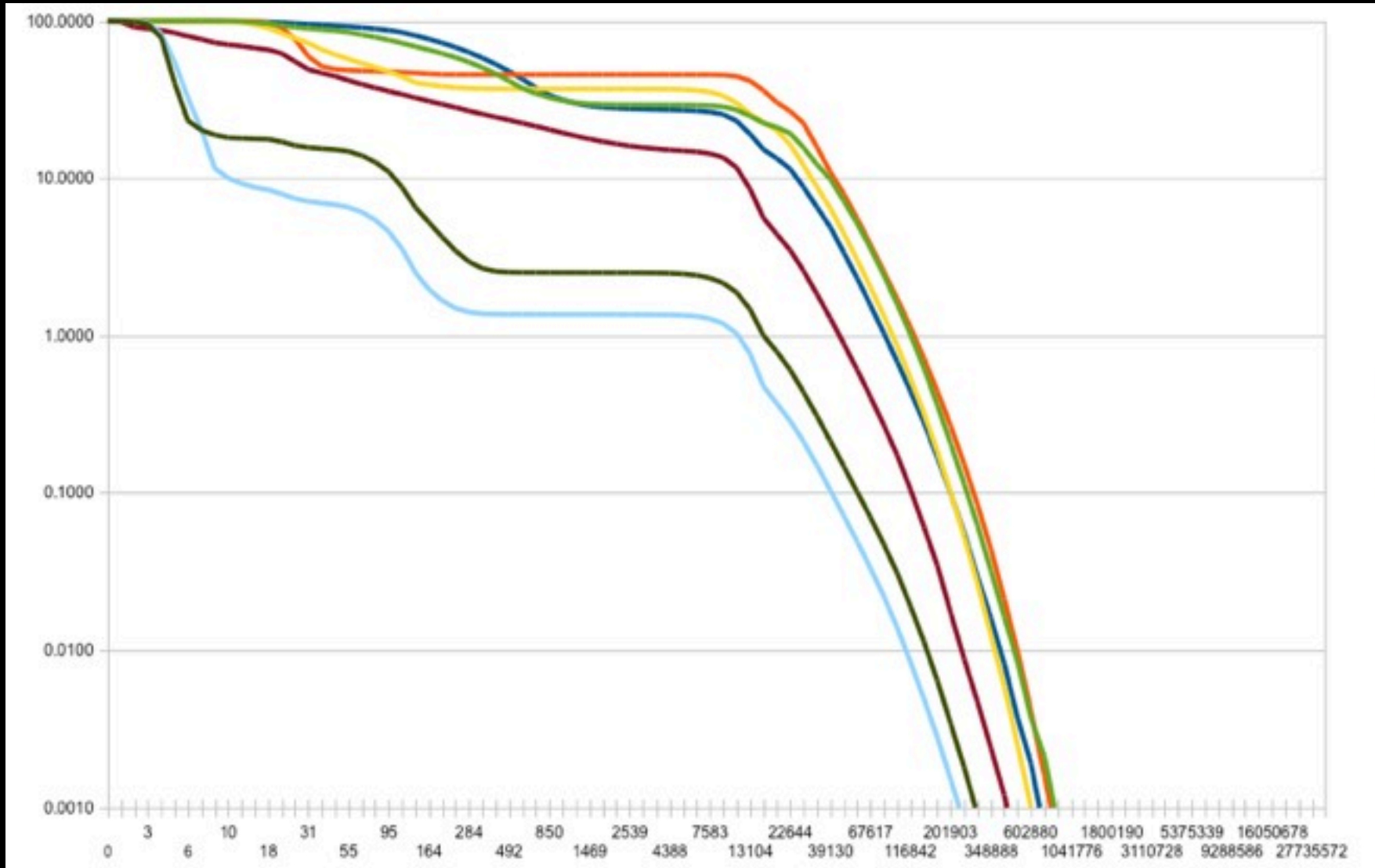
# Latencies

## Column Family



# Read Latencies

Per Column Family



# Misc

- Concurrent Writers
- Switch of Swap
- Quorum has some particularities
  - eventuality still applies