Vitess & CAP
Hitting the sweet spot @YouTube
sougou@google.com
http://ssougou.blogspot.com/
Definitions

- **Consistency**
  - Read-after-write
  - Not same as ACID
  - Opposite of eventual consistency

- **Availability**

- **Partition Tolerance**
  - System is not distributed
  - Distributed: Replicated or Sharded?

- **Per operation**
Connectivity options
CAP partition
Client link failures
Node failure
Motives for replication
Motives for replication

● Durability
● Uptime
● Read scalability
● Latency
Motives for sharding
Motives for sharding

- Write scalability
- Write latency
- Read scalability with Consistency
Summary of concerns

- Consistency
- Availability
- Durability
- Atomicity
- Latency
- Tolerance to various failures
the cake is a lie.
Write options

- Replicated writes
  - Paxos
  - Master-replica scheme
  - Semi-sync replication
  - Multi-master writes

- Sharded writes
  - 2pc
  - Sequenced commits
Read options

- Master reads
- Replica reads
- CAS writes
Corollaries

● Durability, Atomicity or Consistency
  ○ Higher latency
  ○ Availability trade-offs

● Availability
  ○ Lower durability and consistency
  ○ Low latency
Vitess choices

● Single master writes with semi-sync replication
● Sequenced commits for cross-DB transactions
● Master reads for Consistency
● Replica reads for Availability
Crazy ideas

- Paxos master election
- P2P replication