BigTable

A Distributed Storage System for Structured Data.

Fay Chang, Jeffrey Dean, Sanjay Ghemawat, Wilson C. Hsieh, Deborah A. Wallach
Mike Burrows, Tushar Chandra, Andrew Fikes, Robert E. Gruber

Google

- Horizontal Scale
  - Hundreds of servers
  - TB of data (per database)
- Many products
  - (Tools should generalize)

Tradeoff

- **Consistency vs Scale**
  - Stronger Consistency =>
    - Greater Coordination =>
      - (i.e. Paxos and friends)
  - Lower Scale

Scale: GFS

- Unstructured
  - Giant append optimized file system
- Weak Consistency
  - No transactions
  - Weak guarantees write guarantees

Consistency: SQL

- Structured
- Strong Consistency
    - ACID Transactions
- Low Scale
    - One Writer or Sharding

BigTable

- Minimal Coordination
  - Leader election for tablets.
  - Automatic Partitioning
- Minimal Transaction support.
  - Enforced by the tablet server
  - Per row transactions OK

BigTable

- Small Step towards Consistency
  - Better than GFS, but not much
- Good Scale
  - Hundreds of tablet servers
  - Many TB of data

BigTable

- Midpoints on the spectrum exist!
- Among the first NoSQL stores.
  - KV-stores
  - Document Stores
  - Graph Databases
- Partition Data and Don’t Coordinate

RAMP

- Maybe some coordination can scale?
- Read Atomic
- Not Quite ACID
  - Not Serializable

Discussion

- Do other points exist on this spectrum?
- Is there a general data model?
  - Unify K-V vs Document vs Graph
- Is BigTable different from sharded SQL?