

Review: System Types

- File systems
 - GFS/HDFS, NFS...
- Batch processing frameworks
 - Map-reduce, Spark
- Key-value stores
 - Dynamo
- Peer-to-peer
 - Bayou

Review: Algorithms/Components

- Fault-tolerance + Replication
 - Primary-backup (GFS)
 - Consensus (Raft)
 - Sloppy quorum (Dynamo)
 - Lineage (Spark)
- Fail-over
 - Asynchronous safe (Raft)
 - Synchronous assumption (P-B with leases)

Review: Algorithms/Components

- Messaging Semantics
 - RPC
 - At-least once
 - At-most once
- Naming, Routing, Partitioning
 - Explicit map
 - Consistent Hashing

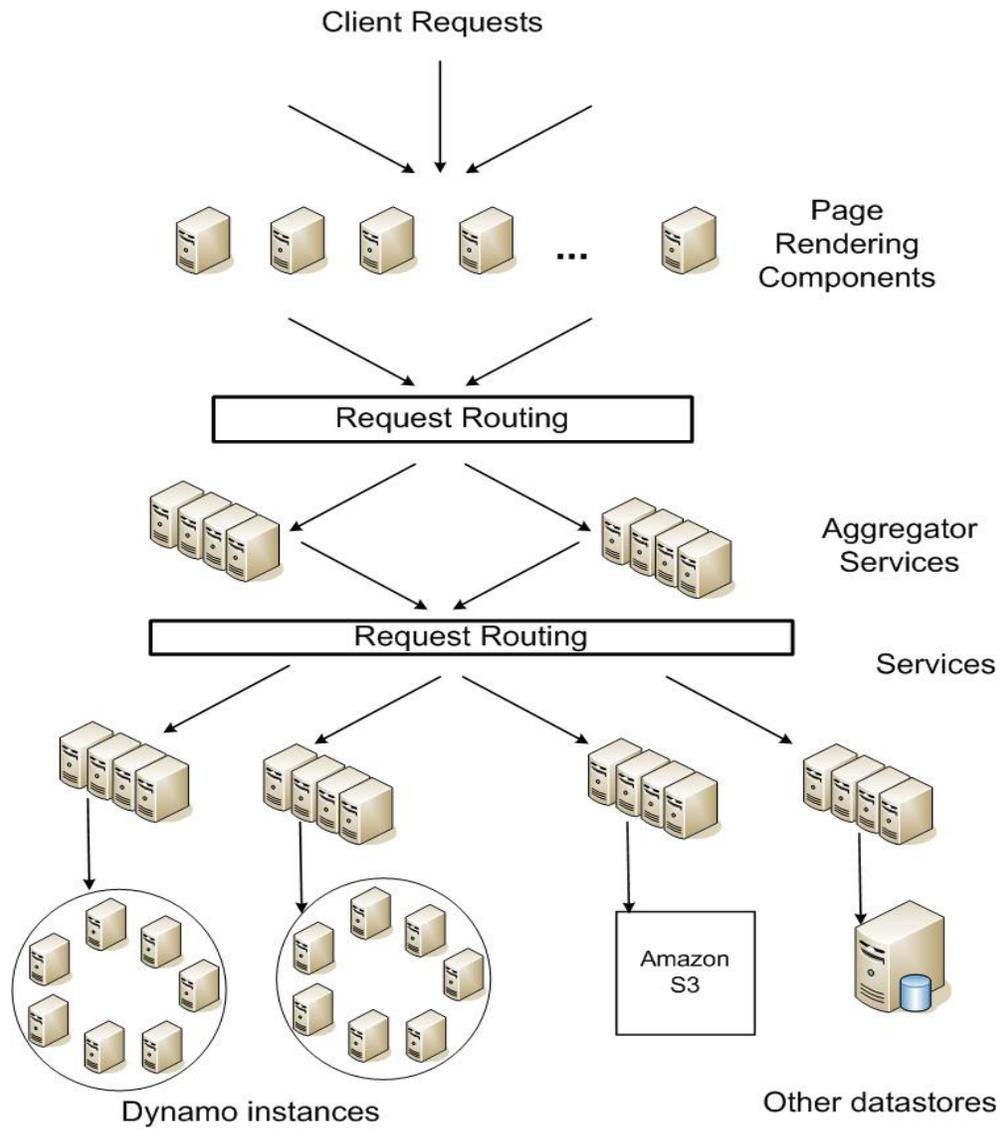
Review: Algorithms/Components

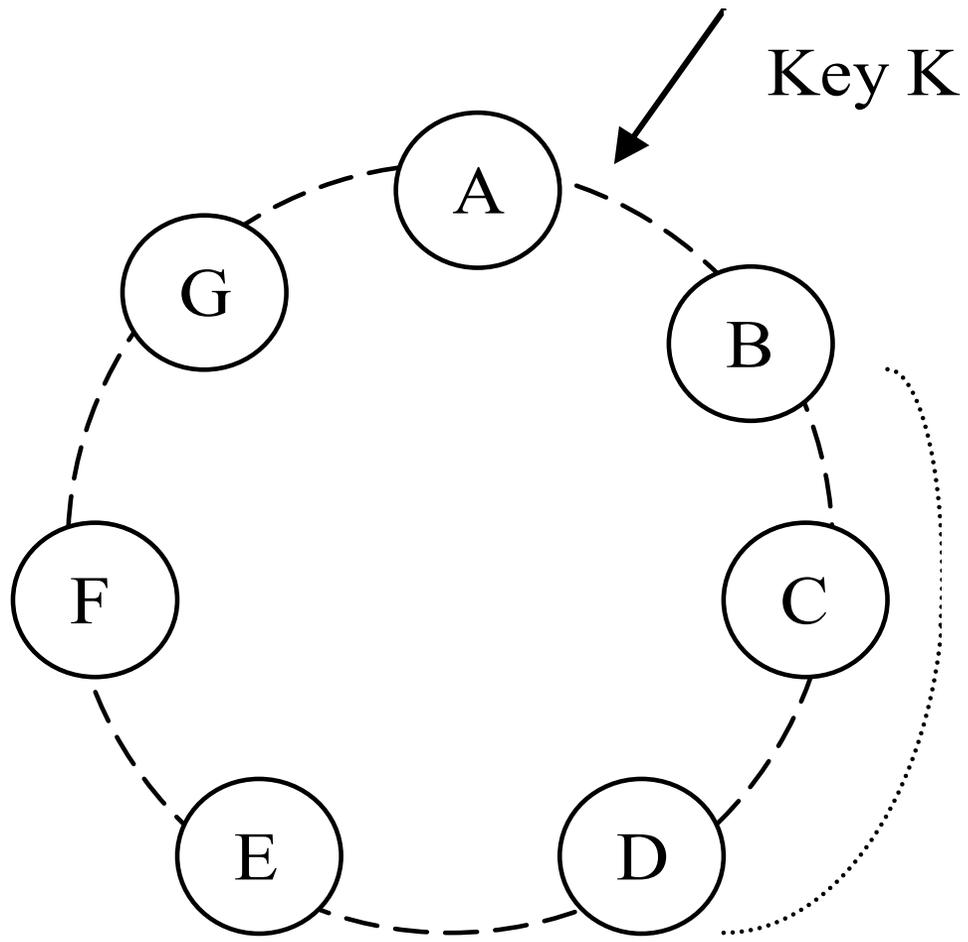
- Clocks
 - Clock drift/clock skew
 - Causality
 - Lamport clocks
 - Vector clocks
- Consistency Models
 - Linearizability
 - Eventual consistency
- CAP Theorem

What's Coming

- Distributed transactions
 - Spanner and more
- Large scale caching infrastructure
 - Facebook
- Cluster Management
 - Borg
- Byzantine Fault Tolerance/P2P
 - PBFT, Bitcoin

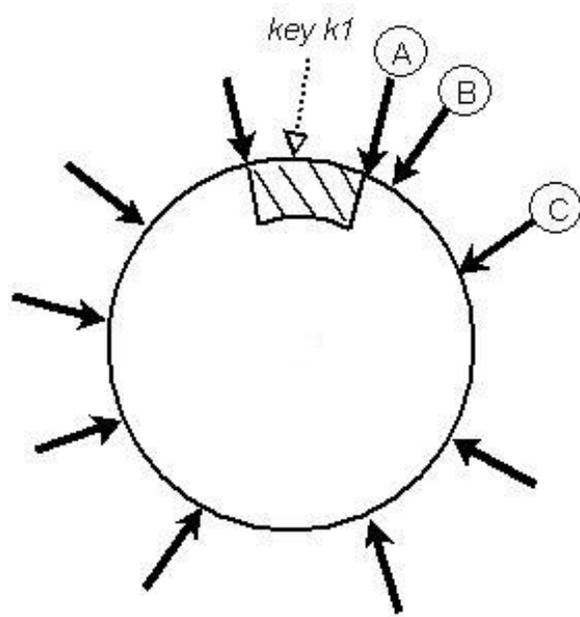
Putting the Pieces Together



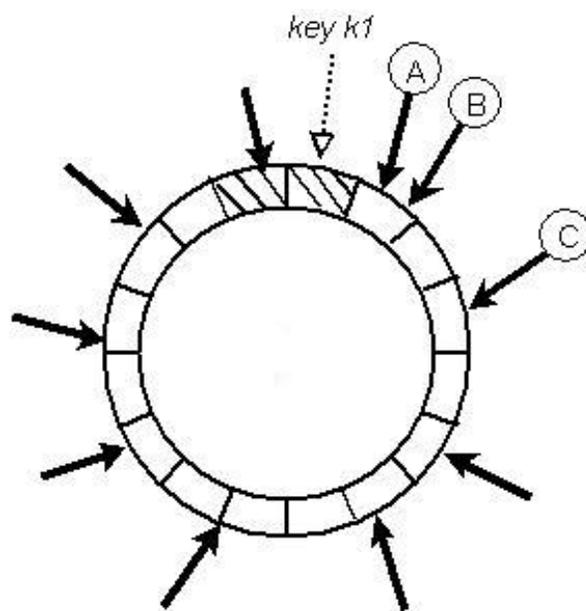


Key K

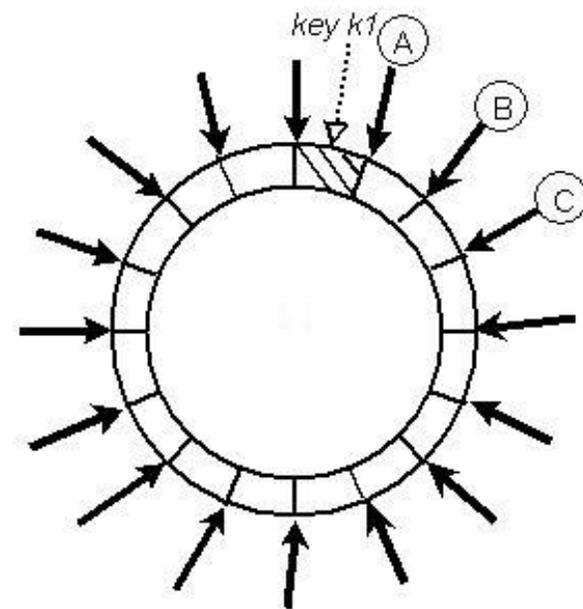
Nodes B, C
and D store
keys in
range (A,B)
including
K.



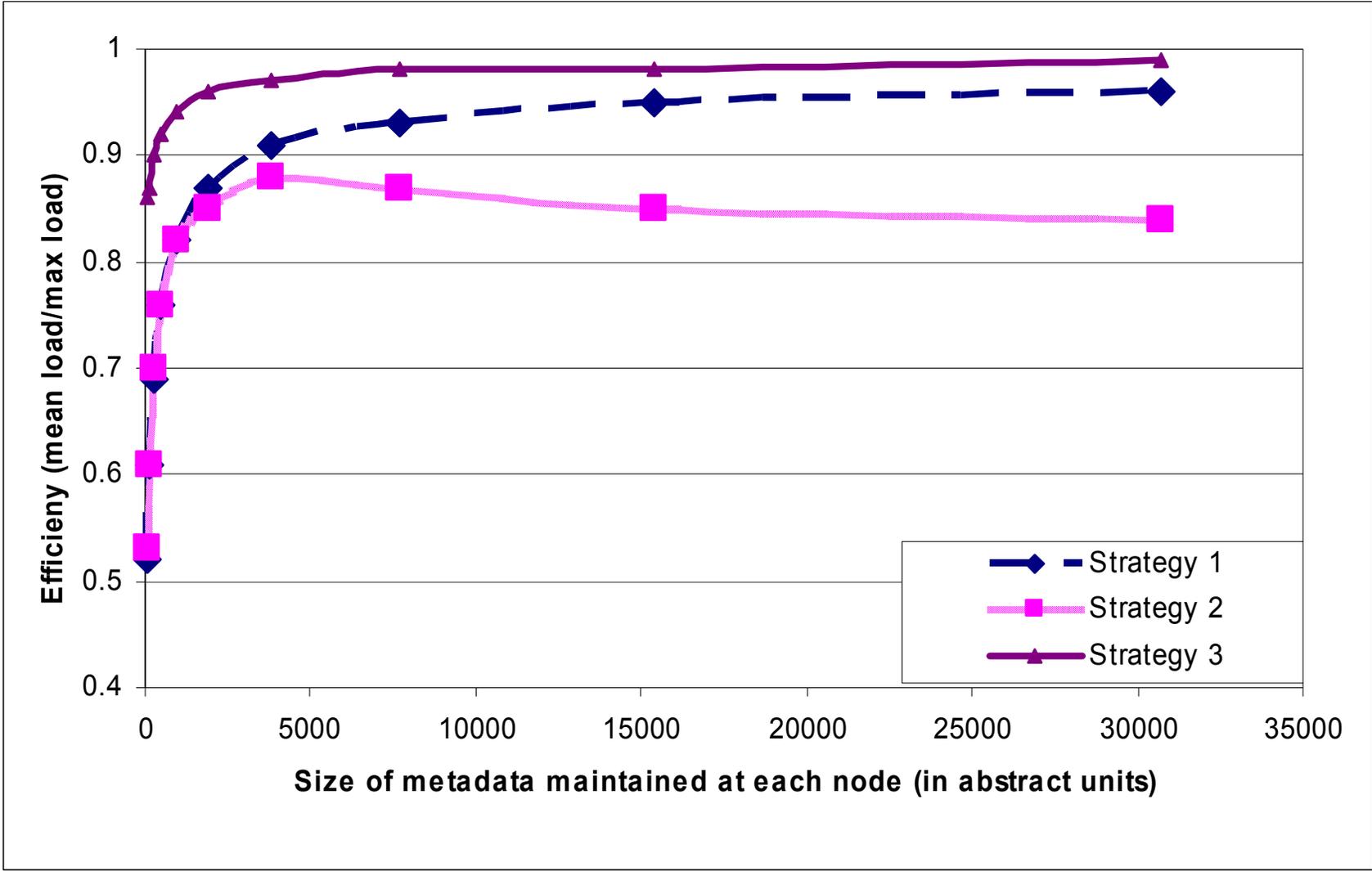
Strategy 1



Strategy 2



Strategy 3



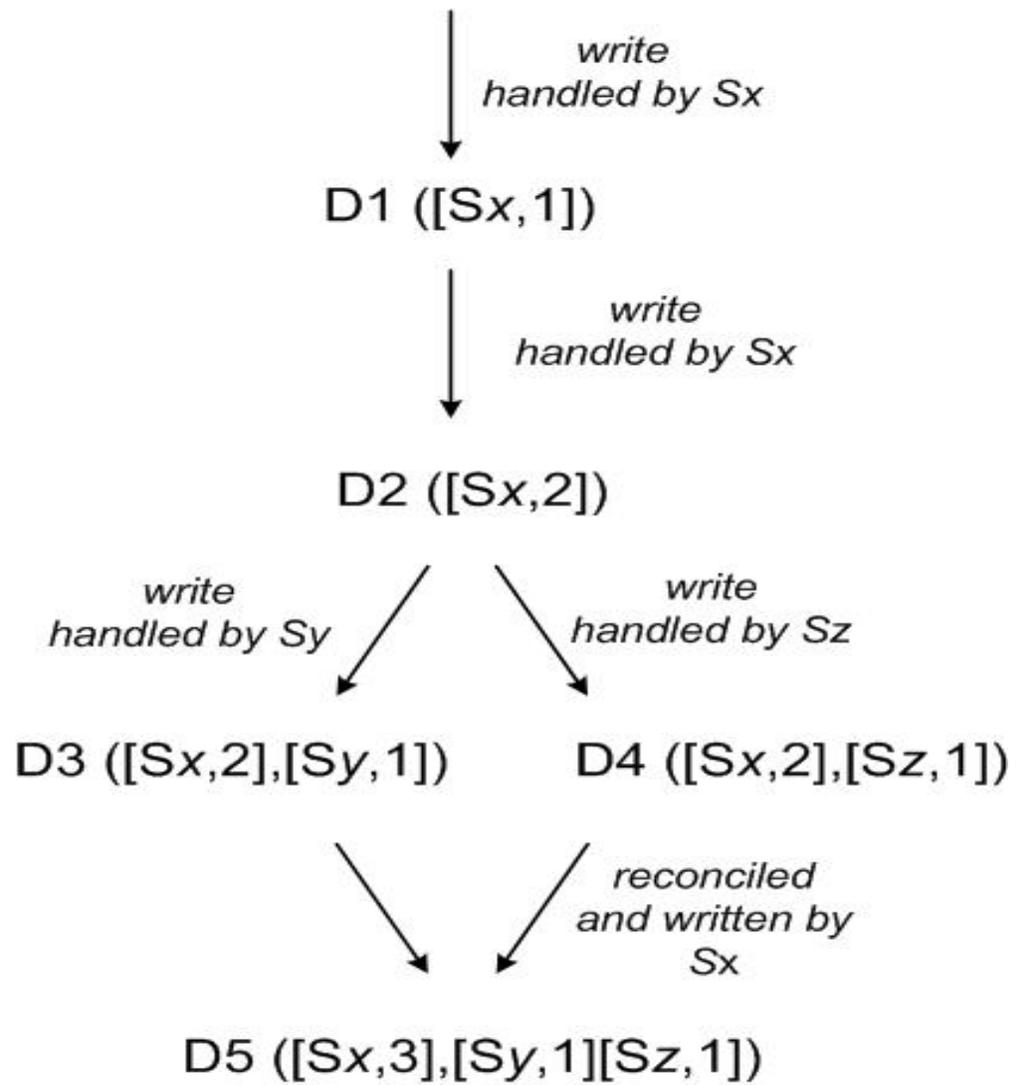


Table 2: Performance of client-driven and server-driven coordination approaches.

	99.9th percentile read latency (ms)	99.9th percentile write latency (ms)	Average read latency (ms)	Average write latency (ms)
Server-driven	68.9	68.5	3.9	4.02
Client-driven	30.4	30.4	1.55	1.9

