Distributed Data Storage and Management Part VI

Saptarshi Pyne
Assistant Professor
Department of Computer Science and Engineering
Indian Institute of Technology Jodhpur, Rajasthan, India 342030

CSL4030 Data Engineering Lectures 16, 17, 18
September 13th, 15th, 18th, 2023

What we discussed in the last class

- Persistent messaging: An alternative model to commit protocols

A few widely used terms

- Primary copy = Master copy = Leader replica
- Replicas = Slaves = Follower replicas

Locking protocols for concurrency control

- Single lock manager protocol
  - Concern: Lock manager is the bottleneck
- Distributed lock manager protocol
  - Concern: Site failures
- Majority Protocol (with ordered lock acquisition)
  - Careful coding is needed

Locking protocols for concurrency control (contd.)

- Biased protocol
  - Prioritizes a ‘shared lock’ on one replica over an ‘exclusive lock’ of all replicas
  - Otherwise, similar to the majority protocol
  - Provides faster read locks (shared) at the cost of slower write locks (exclusive)
  - Useful for data items with much higher demand for read operations than that of write operations, e.g., the IITJ student database

Locking protocols for concurrency control (contd.)

- Quorum consensus protocol (Attiya et al., 1995)
  - All sites are NOT created EQUAL.
    Each **site** is assigned a **nonnegative weight**.
  - **Quorum** = Threshold to be crossed to make a decision, e.g., more than half i.e. majority is a quorum. More than one third could be another quorum.
  - For getting a read lock on data item Q, enough replicas must be locked so that the total weight of the hosting sites is greater than or equal to **read quorum** Q_r.
  - For getting a write lock on data item Q, enough replicas must be locked so that the total weight of the hosting sites is greater than or equal to **write quorum** Q_w.
  - Qr + Qw > S where S = Sum of weights of the sites hosting replicas of Q.
  - 2 * Qw > S
  - Can we simulate majority and biased protocols by setting the site weights and quorums appropriately?

Remaining sub-topics for distributed databases

- Availability
  - High availability at the cost of consistency: The Cloud
- Multi-database systems for heterogeneous distributed databases
- Distributed directory systems for managing data
  - The lightweight directory access protocol (LDAP)

References

- A. SILBERSCHATZ, H.F. KORTH, S. SUDARSHAN (2011), Database System Concepts, McGraw Hill Publications, 6th Edition.
  - Chapter 19. Distributed Databases

- Paper: Bronson et al., “TAO: Facebook’s Distributed Data Store for the Social Graph”, 2013 USENIX Annual Technical Conference (USENIX ATC ‘13).
  - Video: https://www.usenix.org/conference/atc13/technical-sessions/presentation/bronson

Thank you