Distributed Systems

CS425/ECE428

May 3rd 2022

Instructor: Radhika Mittal
Logistics

• Final exam on May 12\textsuperscript{th}, 8-11 am
  • You should be able to see your reservation on PrairieTest.
    • All conflicts and accommodations have been accounted for.
  • Same format as your midterm, but longer:
    • Detailed instructions shared on CampusWire.
  • Unless you have approved accommodations, you have 3 hours to complete the exam from the start time.
• Comprehensive: includes everything covered in the course.
  • Higher weightage assigned to materials that were not covered in midterm syllabus (i.e. Raft and beyond).
PrairieLearn

• Exam format:
  • Multiple choice questions and True/False
    • For questions with multiple choices correct, there is negative marking for selecting incorrect choices to discourage guesswork (the minimum score per question is capped at zero).
  • Numerical questions
    • No step marking!
Exam Syllabus

• All topics covered so far
  • Midterm content
  • Post-midterm content (higher weightage)
    • Starting from Raft, up until distributed datastores.
Exam Syllabus

- **Midterm content (included in finals)**
  - System model and Failures
  - Failure Detection
  - Clock Synchronization
  - Event ordering and Logical Timestamps
  - Global Snapshot
  - Multicast
  - Mutual Exclusion
  - Leader Election
  - Synchronous Consensus and Paxos
Exam Syllabus

• Remaining topics (included in finals)
  • Raft
  • Blockchains
  • Transaction Processing and Concurrency Control
  • Distributed Transactions
  • External consistency and Spanner
  • Distributed Hash Tables
  • MapReduce
  • Distributed Datastores
Disclaimer

• Quick reminder of the relevant concepts we covered in class post midterm.
  • Refer to lecture 15 for a review of pre-midterm topics.

• Not meant to be an exhaustive review!

• Go over the slides for each class.
  • Refer to lecture videos and textbook to fill in gaps in understanding.
Raft

- Algorithm for log consensus. Designed for simplicity.

- What are the guarantees provided by Raft and how?
- How is leader elected?
  - Under what conditions will a process refuse to grant vote?
- What happens when a leader fails or gets disconnected?
- How are log entries appended?
- What leads to missing / extra entries in a server’s log?
- When can log entries be overwritten?
- When can log entries be committed?
Bitcoin / Blockchains

• How is a new transaction added to the log?
  • How is a block mined, and added to a chain?

• What factors determine the rate at which a block is mined?

• What happens if two nodes mine different versions of a block?

• How is information propagated in a Bitcoin network?
Transaction Processing

• What are the ACID properties?
  • How is atomicity achieved?
  • What does consistency mean in this context?
  • What does isolation mean, and how is it achieved?
  • What is durability?
Concurrency Control

• What could go wrong if we don’t have isolation?
  • Lost update problem
  • Inconsistent retrieval problem

• What are conflicting operations?

• What is serial equivalence?

• How can we check if an interleaving is serially equivalent?
Concurrency Control

• Pessimistic Concurrency Control
  • Global lock vs per-object locks vs per-object read/write locks
  • Two-phase locking
  • Deadlocks

• Optimistic Concurrency Control
  • Timestamped ordering
Distributed Transactions

• Meeting ACID requirements for distributed transaction:
  • Two-phase commit for atomicity
  • Distributed deadlock detection with two-phase locking.
External Consistency and Spanner

- High-level design only
  - Maintains synchronized clocks and uses physical timestamps.
    - Reads = snapshot at a physical point in time.
  - What is the concurrency control mechanism used by Spanner?
- Distributed transactions on replicated objects
  - What are the different algorithms / protocols used?
Distributed Hash Tables (Chord)

• What determines the placement of nodes in a Chord ring with m-bit key space?
• Which node is responsible for storing a given key?
• What are the routing table entries maintained by each node:
  • Finger tables
  • r successor entries
• What is the key lookup protocol in Chord?
• How does Chord handle churns?
  • Stabilization protocol.
MapReduce

- Map: creates intermediate key-value pairs
- Reduce: aggregate by key, and run some computation across all values for the key.
- A MapReduce chain comprises of multiple map-reduce pairs.
- Allows easier parallelization.
  - Multiple map/reduce tasks scheduled in parallel across the servers in a cluster.
- Barrier between a map stage and a reduce stage.
  - No reduce task starts before all map tasks are finished.
Distributed Datastores (Cassandra)

- What is CAP theorem?
  - Can only achieve two out of consistency, availability, and partition-tolerance.
- Cassandra: chooses availability, with eventual consistency
  - Key partitioning and replication strategies.
  - How is cluster membership updated?
  - How is a write query executed?
  - How is a read query executed?
  - What are the different consistency levels?
  - What is hinted-handoff and read repair?
Exam Syllabus

• Pre-midterm
  • System model and Failures
  • Failure Detection
  • Clock Synchronization
  • Event ordering and Logical Timestamps
  • Global Snapshot
  • Multicast
  • Mutual Exclusion
  • Leader Election
  • Synchronous Consensus
  • Paxos

• Post-midterm (more weight)
  • Raft
  • Blockchains
  • Transaction Processing and Concurrency Control
  • Distributed Transactions
  • External consistency and Spanner
  • Distributed Hash Tables
  • MapReduce
  • Distributed Datastores
Good luck!