

Lecture 14: Final Review

CSCI4180

Patrick P. C. Lee

Four Areas

- Cloud computing basics
- Cloud computing (MapReduce)
- Cloud storage
- Miscellaneous

Cloud Computing Basics

- Key characteristics
- Service models: SaaS, PaaS, IaaS
- Public/private clouds
- Security issues
- **Understand all terminologies being introduced**

MapReduce and HDFS

- Design features of MapReduce (e.g., map, reduce, combiner, partitioner)
- How does MapReduce handle fault tolerance?
- Design features of HDFS (e.g., namenode, datanode)?
- How does HDFS handle fault tolerance?
- Design features of YARN

MapReduce Programming

- MapReduce programming basics
 - Execution flow
- MapReduce algorithms (design patterns)
 - In-mapper combining, pairs/stripes, order inversion, value-to-key conversion
 - Can you write pseudo-code for a simple MapReduce program?
- Inverted index
- Shortest-path and Pagerank algorithms
- From MapReduce to Spark

Key-Value Stores

- BigTable, HBase
 - Main design features
- How Amazon Dynamo balances performance and availability?
 - Consistent hashing
 - Configurable quorum (N, R, W)

Cloud Storage

- Cost model
- Deduplication
 - Fingerprinting
 - Chunking (Rabin Fingerprinting)
 - Indexing (Bloom Filter, Sparse Indexing, Extreme Binning)
 - Container capping
 - Security
 - Case studies: Dropbox security

Miscellaneous

➤ Facebook

- Haystack: minimizing I/Os in photo search
- f4: minimizing storage overhead for warm BLOBs
- Performance vs. storage efficiency trade-off

➤ Zookeeper

- Coordination mechanism for a distributed system

➤ Tail latency

- Strategies used to tolerate latency

Final Exam

- 2-hour written exam
- Lecture notes, tutorials, assignments
 - Everything except those that I mentioned to be excluded in class
- Two A4 cheat-sheets (both sides)
- Approved calculators allowed
- Make assumptions if needed, and provide justifications

Face-to-face mode:

- **Time: Dec 9, 2022 (Friday), 9:30-11:30am**
- **Venue: Multi-purpose Hall, Chung Chi College**

Make-up Exam Policy

- You must inform the instructor within 12 hours after the original exam time **and** file a formal request to RES
- If you take a make-up exam, **your final exam score will be automatically decremented by 30%.** I hope you can understand this policy, since we need to be fair to the students who work hard to prepare for the regular exam.

Final Exam

➤ How to prepare?

- Understand everything in class notes and assignments
- Not required to read all readings, so long as you understand what the concepts mean

➤ Do past exams

- <http://library.cuhk.edu.hk/>
 - Search by “CSCI4180”
- Ignore questions that we didn't cover

➤ Ask on Piazza

- Please don't ask me about solutions of past exams