Seminar in Databases 2017: Databases in the cloud era
(DRAFT)

Lecturer: Oded Shmueli

Technion Israel Institute of Technology, Haifa 32000, Israel oshmu@cs.technion.ac.il

1 Background - by Lecturer

1. Concurrency Control - Chapter 2 BHG: Serializability Theory
2. Recovery in Distributed Databases - Chapter 7 BHG: Distributed Recovery
3. Replicated Databases - Chapter 8 BHG: Replicated Data

2 Distributed Relational Databases


3 Distributed Storage

2. HBase (Apache) is an open source, non-relational, distributed database modeled after Google's BigTable and written in Java. See Chapter 20, Hadoop book.

4 Performance


5 Concurrency and Recovery


6 Consistency and Properties of Distributed Systems


4. Team of 2


7 Query processing

1. Shumo Chu, Magdalena Balazinska, and Dan Suciu. From theory to practice: Efficient join query evaluation in a parallel database system. In Sellis et al. [24], pages 63–78

2. Kerim Yasin Oktay, Sharad Mehrotra, Vaibhav Khadilkar, and Murat Kantarcioglu. SEMROD: secure and efficient mapreduce over hybrid clouds. In Sellis et al. [24], pages 153–166


4. Mengmeng Liu, Zachary G. Ives, and Boon Thau Loo. Enabling incremental query re-optimization. In Özcan et al. [22], pages 1705–1720

5. Team of 2
   - Immanuel Trummer and Christoph Koch. Multi-objective parametric query optimization. PVLDB, 8(3):221–232, 2014
8 F. Distributed Synchronization Tools

1. Paxos is a family of protocols for solving consensus in a network of unreliable processors. Consensus is the process of agreeing on one result among a group of participants. This problem becomes difficult when the participants or their communication medium may experience failures.

2. Apache ZooKeeper is a software project of the Apache Software Foundation. It is essentially a distributed hierarchical key-value store, which is used to provide a distributed configuration service, synchronization service, and naming registry for large distributed systems.