Readings in Database Systems, 5th Edition: Full Reading List

Background

Joseph M. Hellerstein and Michael Stonebraker. What Goes Around Comes Around. *Readings in Database Systems*, 4th Edition (2005).

Joseph M. Hellerstein, Michael Stonebraker, James Hamilton. Architecture of a Database System. *Foundations and Trends in Databases*, 1, 2 (2007).

Traditional RDBMS Systems

Morton M. Astrahan, Mike W. Blasgen, Donald D. Chamberlin, Kapali P. Eswaran, Jim Gray, Patricia P. Griffiths, W. Frank King III, Raymond A. Lorie, Paul R. McJones, James W. Mehl, Gianfranco R. Putzolu, Irving L. Traiger, Bradford W. Wade, Vera Watson. System R: Relational Approach to Database Management. *ACM Transactions on Database Systems*, 1(2), 1976, 97-137.

Michael Stonebraker and Lawrence A. Rowe. The design of POSTGRES. *SIGMOD*, 1986.

David J. DeWitt, Shahram Ghandeharizadeh, Donovan Schneider, Allan Bricker, Hui-I Hsiao, Rick Rasmussen. The Gamma Database Machine Project. *IEEE Transactions on Knowledge and Data Engineering*, 2(1), 1990, 44-62.

Techniques Everyone Should Know

Patricia G. Selinger, Morton M. Astrahan, Donald D. Chamberlin, Raymond A. Lorie, Thomas G. Price. Access path selection in a relational database management system. *SIGMOD*, 1979.

C. Mohan, Donald J. Haderle, Bruce G. Lindsay, Hamid Pirahesh, Peter M. Schwarz. ARIES: A Transaction Recovery Method Supporting Fine-Granularity Locking and Partial Rollbacks Using Write-Ahead Logging. *ACM Transactions on Database Systems*, 17(1), 1992, 94-162.

Jim Gray, Raymond A. Lorie, Gianfranco R. Putzolu, Irving L. Traiger. Granularity of Locks and Degrees of Consistency in a Shared Data Base. , IBM, September, 1975.

Rakesh Agrawal, Michael J. Carey, Miron Livny. Concurrency Control Performance Modeling: Alternatives and Implications. *ACM Transactions on Database Systems*, 12(4), 1987, 609-654.

C. Mohan, Bruce G. Lindsay, Ron Obermarck. Transaction Management in the R* Distributed Database Management System. *ACM Transactions on Database Systems*, 11(4), 1986, 378-396.

New DBMS Architectures

Mike Stonebraker, Daniel J. Abadi, Adam Batkin, Xuedong Chen, Mitch Cherniack, Miguel Ferreira, Edmond Lau, Amerson Lin, Sam Madden, Elizabeth O'Neil, Pat O'Neil, Alex Rasin, Nga Tran, Stan Zdonik. C-store: A Column-oriented DBMS. *SIGMOD*, 2005.

Cristian Diaconu, Craig Freedman, Erik Ismert, Per-Ake Larson, Pravin Mittal, Ryan Stonecipher, Nitin Verma, Mike Zwilling. Hekaton: SQL Server's Memory-optimized OLTP Engine. *SIGMOD*, 2013.

Stavros Harizopoulos, Daniel J. Abadi, Samuel Madden, Michael Stonebraker. OLTP Through the Looking Glass, and What We Found There. *SIGMOD*, 2008.

Large-Scale Dataflow Engines

Jeff Dean and Sanjay Ghemawat. MapReduce: Simplified Data Processing on Large Clusters. *OSDI*, 2004.

Yuan Yu, Michael Isard, Dennis Fetterly, Mihai Budiu. DryadLINQ: A System for General-Purpose Distributed Data-Parallel Computing Using a High-Level Language. *OSDI*, 2008.

Weak Isolation and Distribution

Atul Adya, Barbara Liskov, and Patrick O'Neil. Generalized Isolation Level Definitions. *ICDE*, 2000.

Giuseppe DeCandia, Deniz Hastorun, Madan Jampani, Gunavardhan Kakulapati, Avinash Lakshman, Alex Pilchin, Swaminathan Sivasubramanian, Peter Vosshall, and Werner Vogels. Dynamo: Amazon's Highly Available Key-Value Store. *SOSP*, 2007.

Eric Brewer. CAP Twelve Years Later: How the "Rules" Have Changed. *IEEE Computer*, 45, 2 (2012).

Query Optimization

Goetz Graefe and William J. McKenna. The Volcano Optimizer Generator: Extensibility and Efficient Search. *ICDE*, 1993.

Ron Avnur and Joseph M. Hellerstein. Eddies: Continuously Adaptive Query Processing. *SIGMOD*, 2000.

Volker Markl, Vijayshankar Raman, David Simmen, Guy Lohman, Hamid Pirahesh, Miso Cilimdzic. Robust Query Processing Through Progressive Approximation. *SIGMOD*, 2004.

Interactive Analytics

Venky Harinarayan, Anand Rajaraman, Jeffrey D. Ullman. Implementing Data Cubes Efficiently. *SIGMOD*, 1996.

Yihong Zhao, Prasad M. Deshpande, Jeffrey F. Naughton. An Array-Based Algorithm for Simultaneous Multidimensional Aggregates. *SIGMOD*, 1997.

Joseph M. Hellerstein, Ron Avnur, Vijayshankar Raman. Informix under CONTROL: Online Query Processing. *Data Mining and Knowledge Discovery*, 4(4), 2000, 281-314.

Sameer Agarwal, Barzan Mozafari, Aurojit Panda, Henry Milner, Samuel Madden, Ion Stoica. BlinkDB: Queries with Bounded Errors and Bounded Response Times on Very Large Data. *EuroSys*, 2013.

Languages

Joachim W. Schmidt. Some High Level Language Constructs for Data of Type Relation. *ACM Transactions on Database Systems*, 2(3), 1977, 247-261.

Arvind Arasu, Shivnath Babu, Jennifer Widom. The CQL Continuous Query Language: Semantic Foundations and Query Execution. *The VLDB Journal*, 15(2), 2006, 121-142.

Peter Alvaro, Neil Conway, Joseph M. Hellerstein, William R. Marczak. Consistency Analysis in Bloom: A CALM and Collected Approach. *CIDR*, 2011.

Web Data

Sergey Brin and Larry Page. The Anatomy of a Large-scale Hypertextual Web Search Engine. *WWW*, 1998.

Eric A. Brewer. Combining Systems and Databases: A Search Engine Retrospective. *Readings in Database Systems, Fourth Edition*, 2005.

Michael J. Cafarella, Alon Halevy, Daisy Zhe Wang, Eugene Wu, Yang Zhang. WebTables: Exploring the Power of Tables on the Web. *VLDB*, 2008.