

21. March 2009

Introduction (Alexis Cheng, Buddy: Daniel Pauli)

The Economist. **Let it rise: A special report on corporate it.** The Economist, 2008.

Brian Hayes. **Cloud computing.** Commun. ACM, 51(7):9-11, 2008.

Ricardo A. Baeza-Yates and Raghu Ramakrishnan. **Data challenges at Yahoo!** In EDBT, pages 652-655, 2008.

Michael J. Cafarella, Edward Chang, Andrew Fikes, Alon Y. Halevy, Wilson C. Hsieh, Alberto Lerner, Jayant Madhavan, and S. Muthukrishnan. **Data management projects at Google.** SIGMOD Record, 37(1):34-38, 2008.

Jeff Dean. **Research Challenges Inspired by Large-Scale Computing at Google.** 2008.

Availability and Consistency (Sacha Krug, Buddy: Fabian Dreier)

Werner Vogels. **Eventually consistent.** Commun. ACM, 52(1):40-44, 2009.

Jim Gray, Pat Helland, Patrick E. O'Neil, and Dennis Shasha. **The dangers of replication and a solution.** In SIGMOD Conference, pages 173-182, 1996.

Additional:

Seth Gilbert and Nancy Lynch. **Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services.** Sigact News, 33:51-59, 2002.

Eric A. Brewer. (invited talk) **Towards robust distributed systems.** In Proceedings of PODC, page 7, 2000.

Time in distributed systems (Samuel Muff, Buddy: ?)

Leslie Lamport. **Time, clocks, and the ordering of events in a distributed system.** Commun. ACM, 21(7):558-565, 1978.

Friedemann Mattern. **Virtual time and global states of distributed systems.** In Cosnard M. et al., editor, Proc. Workshop on Parallel and Distributed Algorithms, pages 215-226, 1989.

Christoph Lenzen, Thomas Locher, and Roger Wattenhofer. **Clock Synchronization with Bounded Global and Local Skew.** Annual IEEE Symposium on Foundations of Computer Science (FOCS), USA, October 2008.

PAXOS (Marcel Schöngens, Buddy : Stephan Merkli)

Leslie Lamport. **Paxos made simple.** In ACM SIGACT News, 2001.

Tushar Deepak Chandra, Robert Griesemer, and Joshua Redstone. **Paxos made live: an engineering perspective.** In Proc. of PODC, pages 398-407, 2007.

Optional:

Leslie Lamport. **Fast paxos.** Distributed Computing, 19(2):79-103, 2006.

Leslie Lamport. **The Part-Time Parliament.** ACM Trans. Comput. Syst., 16(2):133-169, 1998.

Consistent Hashing (Yann Disser, Buddy: Mark Nevill)

David R. Karger, Eric Lehman, Frank Thomson Leighton, Rina Panigrahy, Matthew S. Levine, and Daniel Lewin. **Consistent hashing and random trees: Distributed caching protocols for relieving hot spots on the world wide web.** In STOC, pages 654-663, 1997.

Burton H. Bloom, **Space/time trade-offs in hash coding with allowable errors,** Communications of the ACM, Volume 13, Issue 7, July 1970.

DHT (Benjamin Ernst, Buddy: Stefan Schmid)

Ion Stoica, Robert Morris, David R. Karger, M. Frans Kaashoek, and Hari Balakrishnan. **Chord: A scalable peer-to-peer lookup service for internet applications.** In SIGCOMM, pages 149-160, 2001.

Antony I. T. Rowstron and Peter Druschel. **Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems.** In Middleware, pages 329-350, 2001.

Range Queries (Ruth Syz, Buddy: Conrado Plano)

David R. Karger and Matthias Ruhl. **Simple efficient load-balancing algorithms for peer-to-peer systems.** Theory Comput. Syst., 39(6):787-804, 2006.

Ashwin R. Bharambe, Mukesh Agrawal, and Srinivasan Seshan. **Mercury: supporting scalable multi-attribute range queries.** In SIGCOMM, pages 353-366, 2004.

Karl Aberer, Philippe Cudré -Mauroux, Anwitaman Datta, Zoran Despotovic, Manfred Hauswirth, Magdalena Puceva, and Roman Schmidt. **P-grid: a self-organizing structured p2p system.** SIGMOD Record, 32(3):29-33, 2003.

Gossiping and Anti-Entropy (Tatyana Niolayeva, Buddy: Mike Klausmann)

Robbert van Renesse, Yaron Minsky, and Mark Hayden. **A gossip-based failure detection service.** In Middleware'98, IFIP International Conference on Distributed Systems Platforms and Open Distributed Processing, pages 55-70, England, September 1998.

Karin Petersen, Mike Spreitzer, Douglas B. Terry, Marvin Theimer, and Alan J. Demers. **Flexible update propagation for weakly consistent replication.** In SOSP, pages 288-301, 1997.

Optional:

Ayalvadi J. Ganesh, Anne-Marie Kermarrec, and Laurent Massoulié. **Peer-to-peer membership management for gossip-based protocols.** IEEE Trans. Computers, 52(2):139-149, 2003.

Indranil Gupta, Tushar Deepak Chandra, and Germán S. Goldszmidt. **On scalable and efficient distributed failure detectors.** In PODC, pages 170-179, 2001.

4. April 2009

Design Methodologies (Daniel Pauli, Buddy: ?)

Marcos Kawazoe Aguilera, Arif Merchant, Mehul A. Shah, Alistair C. Veitch, and Christos T. Karamanolis. **Sinfonia: a new paradigm for building scalable distributed systems.** In SOSP, pages 159-174, 2007.

Matt Welsh, David E. Culler, and Eric A. Brewer. **Seda: An architecture for well-conditioned, scalable internet services.** In SOSP, pages 230-243, 2001.

Multi-Database Systems (Fabian Dreier, Buddy: Alexis Cheng)

Christian Plattner and Gustavo Alonso. **Ganymed: Scalable replication for transactional web applications.** In Middleware, pages 155-174, 2004.

Lásaro J. Camargos, Fernando Pedone, and Marcin Wieloch. **Sprint: a middleware for high-performance transaction processing.** In EuroSys, pages 385-398, 2007.

Cloud Locking Systems (?, Buddy: Sacha Krug)

Michael Burrows. **The Chubby Lock Service for Loosely-Coupled Distributed Systems.** In Proc. of OSDI, pages 335-350, 2006.

Apache Software Foundation. **ZooKeeper**, 2008. <http://hadoop.apache.org/zookeeper/>.

Cloud Storage Systems I (Stephan Merkli, Buddy: Samuel Muff)

Sanjay Ghemawat, Howard Gobioff, and Shun-Tak Leung. **The Google File System.** In SOSP, pages 29-43, 2003.

Cloud Storage Systems II (Mark Nevill, Buddy: Marcel Schöngens)

Fay Chang, Jeffrey Dean, Sanjay Ghemawat, Wilson C. Hsieh, Deborah A. Wallach, Michael Burrows, Tushar Chandra, Andrew Fikes, and Robert Gruber. **Bigtable: A distributed storage system for structured data.** In Proc. of OSDI, pages 205-218, 2006.

JJ Furman, Jonas S Karlsson, Jean-Michel Leon, Steve Newman Alex Lloyd, and Philip Zeyliger. **Megastore: A Scalable Data System for User Facing Applications.** Presentation at SIGMOD Products Day, Vancouver, Canada, 2008.

Cloud Storage Systems III (Stefan Schmid, Buddy: Yann Disser)

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.** In Proc. of SOSP, pages 205-220, Oct 2007.

Brian Cooper, Raghu Ramakrishnan, Utkarsh Srivastava, Adam Silberstein, Phil Bohannon, Hans-Arno Jacobsen, Nick Puz, Daniel Weaver, and Ramana Yerneni. **PNUTS: Yahoo!'s hosted data serving platform.** In Proc. of VLDB, 2008.

Cloud Storage Systems IV (Conrado Plan, Buddy: Benjamin Ernst)

Michael Armbrust, Armando Fox, David A. Patterson, Nick Lanham, Beth Trushkowsky, Jesse Trutna, and Haruki Oh. **Scads: Scale independent storage for social computing applications.** 2009.

Alex Sherman, Philip A. Lisiecki, Andy Berkheimer, and Joel Wein. **ACMS: The akamai configuration management system.** In NSDI, 2005.

Other DB-Systems (Mike Klausmann, Buddy: Ruth Syz)

R. Kallman et al. **H-store: a high-performance, distributed main memory transaction processing system.** In Proc. of VLDB, pages 1496-1499, 2008.

M. Stonebraker et al. **The end of an architectural era (it's time for a complete rewrite).** In Proc. of VLDB, pages 1150-1160, 2007.

Map Reduce (? , Buddy: Tatyana Nikolayeva)

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

David DeWitt and Michael Stonebraker. **MapReduce: A Major Step Backward**. The Database Column: <http://www.databasecolumn.com/2008/01/mapreduce-a-major-step-back.html>

David DeWitt and Michael Stonebraker. **MapReduce II**. The Database Column: <http://www.databasecolumn.com/2008/01/mapreduce-continued.html>

(N.B. There is a more comprehensive paper by DeWitt and Stonebraker which has not been published yet. Contact the authors for that.)