APPENDIX A
Bibliography

1. [ABADI12] Abadi, Daniel. 2012. "Consistency Tradeoffs in Modern Distributed Database System Design: CAP is Only Part of the Story." *Computer* 45, no. 2 (February): 37-42. *https://doi.org/10.1109/MC.2012.33*.
2. [ABADI17] Abadi, Daniel. 2017. "Distributed consistency at scale: Spanner vs. Calvin." *Fauna* (blog). April 6, 2017. *https://fauna.com/blog/distributed-consistency-at-scale-spanner-vs-calvin*.
3. [ABADI13] Abadi, Daniel, Peter Boncz, Stavros Harizopoulos, Stratos Idreaos, and Samuel Madden. 2013. *The Design and Implementation of Modern Column-Oriented Database Systems*. Hanover, MA: Now Publishers Inc.
4. [ABRAHAM13] Abraham, Ittai, Danny Dolev, and Joseph Y. Halpern. 2013. "Distributed Protocols for Leader Election: A Game-TheoreticPerspective." In *Distributed Computing*, edited by Yehuda Afek, 61-75. Berlin: Springer, Berlin, Heidelberg.
5. [AGGARWAL88] Aggarwal, Alok, and Jeffrey S. Vitter. 1988. "The input/output complexity of sorting and related problems." *Communications of the ACM* 31, no. 9 (September): 1116-1127. *https://doi.org/10.1145/48529.48535*.
6. [AGRAWAL09] Agrawal, Devesh, Deepak Ganesan, Ramesh Sitaraman, Yanlei Diao, Shashi Singh. 2009. "Lazy-Adaptive Tree: an optimized index structure for flash devices." *Proceedings of the VLDB Endowment* 2, no. 1 (January): 361-372.
7. [AGRAWAL08] Agrawal, Nitin, Vijayan Prabhakaran, Ted Wobber, John D. Davis, Mark Manasse, and Rina Panigrahy. 2008. "Design tradeoffs for SSD performance." *USENIX 2008 Annual Technical Conference (ATC '08)*, 57-70. USENIX.
8. [AGUILERA97] Aguilera, Marcos K., Wei Chen, and Sam Toueg. 1997. "Heartbeat: a Timeout-Free Failure Detector for Quiescent Reliable Communication."

In *Distributed Algorithms*, edited by M. Mavronicolas and P. Tsigas, 126-140. Berlin: Springer, Berlin, Heidelberg.

9. [AGUILERA01] Aguilera, Marcos Kawazoe, Carole Delporte-Gallet, Hugues Fauconnier, and Sam Toueg. 2001. "Stable Leader Election." In *Proceedings of the 15th International Conference on Distributed Computing (DISC '01)*, edited by Jennifer L. Welch, 108-122. London: Springer-Verlag.
10. [AGUILERA16] Aguilera, M. K., and D. B. Terry. 2016. "The Many Faces of Consistency." *Bulletin of the Technical Committee on Data Engineering* 39, no. 1 (March): 3-13.
11. [ALHOUMAILY10] Al-Houmaily, Yousef J. 2010. "Atomic commit protocols, their integration, and their optimisations in distributed database systems." *International Journal of Intelligent Information and Database Systems* 4, no. 4 (September): 373–412. *https://doi.org/10.1504/IJIIDS.2010.035582*.
12. [ARJOMANDI83] Arjomandi, Eshrat, Michael J. Fischer, and Nancy A. Lynch. 1983. "Efficiency of Synchronous Versus Asynchronous Distributed Systems." *Journal of the ACM* 30, no. 3 (July): 449-456. *https://doi.org/10.1145/2402.322387*.
13. [ARULRAJ17] Arulraj, J. and A. Pavlo. 2017. "How to Build a Non-Volatile Memory Database Management System." In *Proceedings of the 2017 ACM International Conference on Management of Data*: 1753-1758. *https://doi.org/10.1145/3035918.3054780*.
14. [ATHANASSOULIS16] Athanassoulis, Manos, Michael S. Kester, Lukas M. Maas, Radu Stoica, Stratos Idreos, Anastasia Ailamaki, and Mark Callaghan. 2016. "Designing Access Methods: The RUM Conjecture." In *International Conference on Extending Database Technology (EDBT)*. *https://stratos.seas.harvard.edu/files/stratos/files/rum.pdf*.
15. [ATTIYA94] Attiyaand, Hagit and Jennifer L. Welch. 1994. "Sequential consistency versus linearizability." *ACM Transactions on Computer Systems* 12, no. 2 (May): 91-122. *https://doi.org/10.1145/176575.176576*.
16. [BABAOGLU93] Babaoglu, Ozalp and Sam Toueg. 1993. "Understanding Non-Blocking Atomic Commitment." Technical Report. University of Bologna.
17. [BAILIS14a] Bailis, Peter. 2014. "Linearizability versus Serializability." *Highly Available, Seldom Consistent* (blog). September 24, 2014. *https://www.bailis.org/blog/linearizability-versus-serializability*.
18. [BAILIS14b] Bailis, Peter, Alan Fekete, Michael J. Franklin, Ali Ghodsi, Joseph M. Hellerstein, and Ion Stoica. 2014. "Coordination Avoidance in Database Systems." *Proceedings of the VLDB Endowment* 8, no. 3 (November): 185-196. *https://doi.org/10.14778/2735508.2735509*.

19. [BAILIS14c] Bailis, Peter, Alan Fekete, Ali Ghodsi, Joseph M. Hellerstein, and Ion Stoica. 2014. "Scalable Atomic Visibility with RAMP Transactions." *ACM Transactions on Database Systems* 41, no. 3 (July). *https://doi.org/10.1145/2909870*.

20. [BARTLETT16] Bartlett, Robert P. III, and Justin McCrary. 2016. "How Rigged Are Stock Markets?: Evidence From Microsecond Timestamps." UC Berkeley Public Law Research Paper. *https://doi.org/10.2139/ssrn.2812123*.

21. [BAUDET19] Baudet, Mathieu, Avery Ching, Andrey Chursin, George Danezis, François Garillot, Zekun Li, Dahlia Malkhi, Oded Naor, Dmitri Perelman, and Alberto Sonnino. 2019. "State Machine Replication in the Libra Blockchain." *https://developers.libra.org/docs/assets/papers/libra-consensus-state-machine-replication-in-the-libra-blockchain.pdf*.

22. [BAYER72] Bayer, R., and E. M. McCreight. 1972. "Organization and maintenance of large ordered indices." *Acta Informatica* 1, no. 3 (September): 173-189. *https://doi.org/10.1007/BF00288683*.

23. [BEDALY69] Belady, L. A., R. A. Nelson, and G. S. Shedler. 1969. "An anomaly in space-time characteristics of certain programs running in a paging machine." *Communications of the ACM* 12, no. 6 (June): 349-353. *https://doi.org/10.1145/363011.363155*.

24. [BENDER05] Bender, Michael A., Erik D. Demaine, and Martin Farach-Colton. 2005. "Cache-Oblivious B-Trees." *SIAM Journal on Computing* 35, no. 2 (August): 341-358. *https://doi.org/10.1137/S0097539701389956*.

25. [BERENSON95] Berenson, Hal, Phil Bernstein, Jim Gray, Jim Melton, Elizabeth O'Neil, and Patrick O'Neil. 1995. "A critique of ANSI SQL isolation levels." *ACM SIGMOD Record* 24, no. 2 (May): 1-10. *https://doi.org/10.1145/568271.223785*.

26. [BERNSTEIN87] Bernstein, Philip A., Vassco Hadzilacos, and Nathan Goodman. 1987. *Concurrency Control and Recovery in Database Systems*. Boston: Addison-Wesley Longman.

27. [BERNSTEIN09] Bernstein, Philip A. and Eric Newcomer. 2009. *Principles of Transaction Processing*. San Francisco: Morgan Kaufmann.

28. [BHATTACHARJEE17] Bhattacharjee, Abhishek, Daniel Lustig, and Margaret Martonosi. 2017. *Architectural and Operating System Support for Virtual Memory*. San Rafael, CA: Morgan & Claypool Publishers.

29. [BIRMAN07] Birman, Ken. 2007. "The promise, and limitations, of gossip protocols." *ACM SIGOPS Operating Systems Review* 41, no. 5 (October): 8-13. *https://doi.org/10.1145/1317379.1317382*.

30. [BIRMAN10] Birman, Ken. 2010. "A History of the Virtual Synchrony Replication Model" In *Replication*, edited by Bernadette Charron-Bost, Fernando Pedone, and André Schiper, 91-120. Berlin: Springer-Verlag, Berlin, Heidelberg.

31. [BIRMAN06] Birman, Ken, Coimbatore Chandersekaran, Danny Dolev, Robbert vanRenesse. 2006. "How the Hidden Hand Shapes the Market for Software Reliability." In *First Workshop on Applied Software Reliability (WASR 2006)*. IEEE.

32. [BIYIKOGLU13] Biyikoglu, Cihan. 2013. "Under the Hood: Redis CRDTs (Conflict-free Replicated Data Types)." http://lp.redislabs.com/rs/915-NFD-128/images/WP-RedisLabs-Redis-Conflict-free-Replicated-Data-Types.pdf.

33. [BJØRLING17] Bjørling, Matias, Javier González, and Philippe Bonnet. 2017. "LightNVM: the Linux open-channel SSD subsystem." In *Proceedings of the 15th Usenix Conference on File and Storage Technologies (FAST'17)*, 359-373. USENIX.

34. [BLOOM70] Bloom, Burton H. 1970. "Space/time trade-offs in hash coding with allowable errors." *Communications of the ACM* 13, no. 7 (July): 422-426. *https://doi.org/10.1145/362686.362692*.

35. [BREWER00] Brewer, Eric. 2000. "Towards robust distributed systems." *Proceedings of the nineteenth annual ACM symposium on Principles of distributed computing (PODC '00)*. New York: Association for Computing Machinery. *https://doi.org/10.1145/343477.343502*.

36. [BRZEZINSKI03] Brzezinski, Jerzy, Cezary Sobaniec, and Dariusz Wawrzyniak. 2003. "Session Guarantees to Achieve PRAM Consistency of Replicated Shared Objects." In *Parallel Processing and Applied Mathematics*, 1–8. Berlin: Springer, Berlin, Heidelberg.

37. [BUCHMAN18] Buchman, Ethan, Jae Kwon, and Zarko Milosevic. 2018. "The latest gossip on BFT consensus." *https://arxiv.org/pdf/1807.04938.pdf*.

38. [CACHIN11] Cachin, Christian, Rachid Guerraoui, and Luis Rodrigues. 2011. *Introduction to Reliable and Secure Distributed Programming (2nd Ed.)*. New York: Springer.

39. [CASTRO99] Castro, Miguel. and Barbara Liskov. 1999. "Practical Byzantine Fault Tolerance." In *OSDI '99 Proceedings of the third symposium on Operating systems design and implementation*, 173-186.

40. [CESATI05] Cesati, Marco, and Daniel P. Bovet. 2005. *Understanding the Linux Kernel*. Third Edition. Sebastopol: O'Reilly Media, Inc.

41. [CHAMBERLIN81] Chamberlin, Donald D., Morton M. Astrahan, Michael W. Blasgen, James N. Gray, W. Frank King, Bruce G. Lindsay, Raymond Lorie, James W. Mehl, Thomas G. Price, Franco Putzolu, Patricia Griffiths Selinger, Mario Schkolnick, Donald R. Slutz, Irving L. Traiger, Bradford W. Wade, and Robert A. Yost. 1981. "A history and evaluation of System R." *Communications of the ACM* 24, no. 10 (October): 632–646. *https://doi.org/10.1145/358769.358784*.

42. [CHANDRA07] Chandra, Tushar D., Robert Griesemer, and Joshua Redstone. 2007. "Paxos made live: an engineering perspective." In *Proceedings of the twenty-sixth annual ACM symposium on Principles of distributed computing (PODC '07)*,

398-407. New York: Association for Computing Machinery. *https://doi.org/10.1145/1281100.1281103*.

43. [CHANDRA96] Chandra, Tushar Deepak, and Sam Toueg. 1996. "Unreliable failure detectors for reliable distributed systems." *Journal of the ACM* 43, no. 2 (March): 225-267. *https://doi.org/10.1145/226643.226647*.

44. [CHANG79] Chang, Ernest, and Rosemary Roberts. 1979. "An improved algorithm for decentralized extrema-finding in circular configurations of processes." *Communications of the ACM* 22, no. 5 (May): 281–283. *https://doi.org/10.1145/359104.359108*.

45. [CHANG06] Chang, Fay, Jeffrey Dean, Sanjay Ghemawat, Wilson C. Hsieh, Deborah A.Wallach, Mike Burrows, Tushar Chandra, Andrew Fikes, and Robert E. Gruber. 2006. "Bigtable: A Distributed Storage System for Structured Data." In *7th USENIX Symposium on Operating Systems Design and Implementation (OSDI '06)*. USENIX.

46. [CHAZELLE86] Chazelle, Bernard, and Leonidas J. Guibas. 1986. "Fractional Cascading, A Data Structuring Technique." *Algorithmica* 1: 133-162. *https://doi.org/10.1007/BF01840440*.

47. [CHOCKLER15] Chockler, Gregory, and Dahlia Malkhi. 2015. "Active disk paxos with infinitely many processes." In *Proceedings of the twenty-first annual symposium on Principles of distributed computing (PODC '02)*, 78-87. New York: Association for Computing Machinery. *https://doi.org/10.1145/571825.571837*.

48. [COMER79] Comer, Douglas. 1979. "Ubiquitous B-Tree." *ACM Computing Survey* 11, no. 2 (June): 121-137. *https://doi.org/10.1145/356770.356776*.

49. [CORBET18] Corbet, Jonathan. 2018. "PostgreSQL's fsync() surprise." *https://lwn.net/Articles/752063*.

50. [CORBETT12] Corbett, James C., Jeffrey Dean, Andrew Fikes, Christopher Frost, JJ Furman, Sanjay Ghemawat, Andrey Gubarev, Christopher Heiser, Peter Hochschild, Wilson Hsieh, Sebastian Kanthak, Eugene Kogan, Hongyi Li, Alexander Lloyd, Sergey Melnik, David Mwaura, David Nagle, Sean Quinlan, Rajesh Rao, Lindsay Rolig, Yasushi Saito, Michal Szymaniak, Christopher Taylor, Ruth Wang, and Dale Woodford. 2012. "Spanner: Google's Globally-Distributed Database." In *10th USENIX Symposium on Operating Systems Design and Implementation (OSDI '12)*, 261-264. USENIX.

51. [CORMODE04] Cormode, G. and S. Muthukrishnan. 2004. "An improved data stream summary: The count-min sketch and its applications." *Journal of Algorithms* 55, No. 1 (April): 58-75. *https://doi.org/10.1016/j.jalgor.2003.12.001*.

52. [CORMODE11] Cormode, Graham, and S. Muthukrishnan. 2011. "Approximating Data with the Count-Min Data Structure." *http://dimacs.rutgers.edu/~graham/pubs/papers/cmsoft.pdf*.

53. [CORMODE12] Cormode, Graham and Senthilmurugan Muthukrishnan. 2012. "Approximating Data with the Count-Min Data Structure."

54. [CHRISTIAN91] Cristian, Flavin. 1991. "Understanding fault-tolerant distributed systems." *Communications of the ACM* 34, no. 2 (February): 56-78. *https://doi.org/10.1145/102792.102801*.

55. [DAILY13] Daily, John. 2013. "Clocks Are Bad, Or, Welcome to the Wonderful World of Distributed Systems." *Riak* (blog). November 12, 2013. *https://riak.com/clocks-are-bad-or-welcome-to-distributed-systems*.

56. [DECANDIA07] DeCandia, Giuseppe, Deniz Hastorun, Madan Jampani, Gunavardhan Kakulapati, Avinash Lakshman, Alex Pilchin, Swaminathan Sivasubramanian, Peter Vosshall, and Werner Vogels. 2007. "Dynamo: amazon's highly available key-value store." *SIGOPS Operating Systems Review* 41, no. 6 (October): 205-220. *https://doi.org/10.1145/1323293.1294281*.

57. [DECHEV10] Dechev, Damian, Peter Pirkelbauer, and Bjarne Stroustrup. 2010. "Understanding and Effectively Preventing the ABA Problem in Descriptor-Based Lock-Free Designs." *Proceedings of the 2010 13th IEEE International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing (ISORC '10)*: 185–192. *https://doi.org/10.1109/ISORC.2010.10*.

58. [DEMAINE02] Demaine, Erik D. 2002. "Cache-Oblivious Algorithms and Data Structures." In *Lecture Notes from the EEF Summer School on Massive Data Sets*. Denmark: University of Aarhus.

59. [DEMERS87] Demers, Alan, Dan Greene, Carl Hauser, Wes Irish, John Larson, Scott Shenker, Howard Sturgis, Dan Swinehart, and Doug Terry. 1987. "Epidemic algorithms for replicated database maintenance." In *Proceedings of the sixth annual ACM Symposium on Principles of distributed computing (PODC '87)*, 1-12. New York: Association for Computing Machinery. *https://doi.org/10.1145/41840.41841*.

60. [DENNING68] Denning, Peter J. 1968. "The working set model for program behavior". *Communications of the ACM* 11, no. 5 (May): 323-333. *https://doi.org/10.1145/363095.363141*.

61. [DIACONU13] Diaconu, Cristian, Craig Freedman, Erik Ismert, Per-Åke Larson, Pravin Mittal, Ryan Stonecipher, Nitin Verma, and Mike Zwilling. 2013. "Hekaton: SQL Server's Memory-Optimized OLTP Engine." In *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data (SIGMOD '13)*, 1243-1254. New York: Association for Computing Machinery. *https://doi.org/10.1145/2463676.2463710*.

62. [DOWNEY12] Downey, Jim. 2012. "Be Careful with Sloppy Quorums." *Jim Downey* (blog). March 5, 2012. *https://jimdowney.net/2012/03/05/be-careful-with-sloppy-quorums*.

63. [DREPPER07] Drepper, Ulrich. 2007. *What Every Programmer Should Know About Memory*. Boston: Red Hat, Inc.

64. [DUNAGAN04] Dunagan, John, Nicholas J. A. Harvey, Michael B. Jones, Dejan Kostić, MarvinTheimer, and Alec Wolman. 2004. "FUSE: lightweight guaranteed distributed failure notification." In *Proceedings of the 6th conference on Symposium on Operating Systems Design & Implementation - Volume 6 (OSDI'04)*, 11-11. USENIX.

65. [DWORK88] Dwork, Cynthia, Nancy Lynch, and Larry Stockmeyer. 1988. "Consensus in the presence of partial synchrony." *Journal of the ACM* 35, no. 2 (April): 288-323. *https://doi.org/10.1145/42282.42283*.

66. [EINZIGER15] Einziger, Gil and Roy Friedman. 2015. "A formal analysis of conservative update based approximate counting." In *2015 International Conference on Computing, Networking and Communications (ICNC)*, 260-264. IEEE.

67. [EINZIGER17] Einziger, Gil, Roy Friedman, and Ben Manes. 2017. "TinyLFU: A Highly Efficient Cache Admission Policy." In *2014 22nd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing*, 146-153. IEEE.

68. [ELLIS11] Ellis, Jonathan. 2011. "Understanding Hinted Handoff." *Datastax* (blog). May 31, 2011. *https://www.datastax.com/dev/blog/understanding-hinted-handoff*.

69. [ELLIS13] Ellis, Jonathan. 2013. "Why Cassandra doesn't need vector clocks." *Datastax* (blog). September 3, 2013. *https://www.datastax.com/dev/blog/why-cassandra-doesnt-need-vector-clocks*.

70. [ELMASRI11] Elmasri, Ramez and Shamkant Navathe. 2011. *Fundamentals of Database Systems (6th Ed.)*. Boston: Pearson.

71. [FEKETE04] Fekete, Alan, Elizabeth O'Neil, and Patrick O'Neil. 2004. "A read-only transaction anomaly under snapshot isolation." *ACM SIGMOD Record* 33, no. 3 (September): 12-14. *https://doi.org/10.1145/1031570.1031573*.

72. [FISCHER85] Fischer, Michael J., Nancy A. Lynch, and Michael S. Paterson. 1985. "Impossibility of distributed consensus with one faulty process." *Journal of the ACM* 32, 2 (April): 374-382. *https://doi.org/10.1145/3149.214121*.

73. [FLAJOLET12] Flajolet, Philippe, Eric Fusy, Olivier Gandouet, and Frédéric Meunier. 2012. "HyperLogLog: The analysis of a near-optimal cardinality estimation algorithm." In *AOFA '07: Proceedings of the 2007 International Conference on Analysis of Algorithms*.

74. [FOWLER11] Fowler, Martin. 2011. "The LMAX Architecture." *Martin Fowler*. July 12, 2011. *https://martinfowler.com/articles/lmax.html*.

75. [FOX99] Fox, Armando and Eric A. Brewer. 1999. "Harvest, Yield, and Scalable Tolerant Systems." In *Proceedings of the Seventh Workshop on Hot Topics in Operating Systems*, 174-178.

76. [FREILING11] Freiling, Felix C., Rachid Guerraoui, and Petr Kuznetsov. 2011. "The failure detector abstraction." *ACM Computing Surveys* 43, no. 2 (January): Article 9. *https://doi.org/10.1145/1883612.1883616*.

77. [MOLINA82] Garcia-Molina, H. 1982. "Elections in a Distributed Computing System." *IEEE Transactions on Computers* 31, no. 1 (January): 48-59. *https://dx.doi.org/10.1109/TC.1982.1675885*.

78. [MOLINA92] Garcia-Molina, H. and K. Salem. 1992. "Main Memory Database Systems: An Overview." *IEEE Transactions on Knowledge and Data Engineering* 4, no. 6 (December): 509-516. *https://doi.org/10.1109/69.180602*.

79. [MOLINA08] Garcia-Molina, Hector, Jeffrey D. Ullman, and Jennifer Widom. 2008. *Database Systems: The Complete Book* (2nd Ed.). Boston: Pearson.

80. [GEORGOPOULOS16] Georgopoulos, Georgios. 2016. "Memory Consistency Models of Modern CPUs." *https://es.cs.uni-kl.de/publications/datarsg/Geor16.pdf*.

81. [GHOLIPOUR09] Gholipour, Majid, M. S. Kordafshari, Mohsen Jahanshahi, and Amir Masoud Rahmani. 2009. "A New Approach For Election Algorithm in Distributed Systems." In *2009 Second International Conference on Communication Theory, Reliability, and Quality of Service*, 70-74. IEEE. *https://doi.org/10.1109/CTRQ.2009.32*.

82. [GIAMPAOLO98] Giampaolo, Dominic. 1998. *Practical File System Design with the be File System*. San Francisco: Morgan Kaufmann.

83. [GILAD17] Gilad, Yossi, Rotem Hemo, Silvio Micali, Georgios Vlachos, and Nickolai Zeldovich. 2017. "Algorand: Scaling Byzantine Agreements for Cryptocurrencies." *Proceedings of the 26th Symposium on Operating Systems Principles* (October): 51–68. *https://doi.org/10.1145/3132747.3132757*.

84. [GILBERT02] Gilbert, Seth and Nancy Lynch. 2002. "Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services." *ACM SIGACT News* 33, no. 2 (June): 51-59. *https://doi.org/10.1145/564585.564601*.

85. [GILBERT12] Gilbert, Seth and Nancy Lynch. 2012. "Perspectives on the CAP Theorem." *Computer* 45, no. 2 (February): 30-36. *https://doi.org/10.1109/MC.2011.389*.

86. [GOMES17] Gomes, Victor B. F., Martin Kleppmann, Dominic P. Mulligan, and Alastair R. Beresford. 2017. "Verifying strong eventual consistency in distributed systems." *Proceedings of the ACM on Programming Languages* 1 (October). *https://doi.org/10.1145/3133933*.

87. [GONÇALVES15] Gonçalves, Ricardo, Paulo Sérgio Almeida, Carlos Baquero, and Victor Fonte. 2015. "Concise Server-Wide Causality Management for Eventually Consistent Data Stores." In *Distributed Applications and Interoperable Systems*, 66-79. Berlin: Springer.

88. [GOOSSAERT14] Goossaert, Emmanuel. 2014. "Coding For SSDs." *CodeCapsule* (blog). February 12, 2014. *http://codecapsule.com/2014/02/12/coding-for-ssds-part-1-introduction-and-table-of-contents*.

89. [GRAEFE04] Graefe, Goetz. 2004. "Write-Optimized B-Trees." In *Proceedings of the Thirtieth international conference on Very large data bases - Volume 30 (VLDB '04)*, 672-683. VLDB Endowment.

90. [GRAEFE07] Graefe, Goetz. 2007. "Hierarchical locking in B-tree indexes." *https://www.semanticscholar.org/paper/Hierarchical-locking-in-B-tree-indexes-Graefe/270669b1eb0d31a99fe99bec67e47e9b11b4553f*.

91. [GRAEFE10] Graefe, Goetz. 2010. "A survey of B-tree locking techniques." *ACM Transactions on Database Systems* 35, no. 3, (July). *https://doi.org/10.1145/1806907.1806908*.

92. [GRAEFE11] Graefe, Goetz. 2011. "Modern B-Tree Techniques." *Foundations and Trends in Databases* 3, no. 4 (April): 203-402. *https://doi.org/10.1561/1900000028*.

93. [GRAY05] Gray, Jim, and Catharine van Ingen. 2005. "Empirical Measurements of Disk Failure Rates and Error Rates." Accessed March 4, 2013. *https://arxiv.org/pdf/cs/0701166.pdf*.

94. [GRAY04] Gray, Jim, and Leslie Lamport. 2004. "Consensus on Transaction Commit." *ACM Transactions on Database Systems* 31, no. 1 (March): 133-160. *https://doi.org/10.1145/1132863.1132867*.

95. [GUERRAOUI07] Guerraoui, Rachid. 2007. "Revisiting the relationship between non-blocking atomiccommitment and consensus." In *Distributed Algorithms*, 87-100. Berlin: Springer, Berlin, Heidelberg. *https://doi.org/10.1007/BFb0022140*.

96. [GUERRAOUI97] Guerraoui, Rachid, and André Schiper. 1997. "Consensus: The Big Misunderstanding." In *Proceedings of the Sixth IEEE Computer Society Workshop on Future Trends of Distributed Computing Systems*, 183-188. IEEE.

97. [GUPTA01] Gupta, Indranil, Tushar D. Chandra, and Germán S. Goldszmidt. 2001. "On scalable and efficient distributed failure detectors." In *Proceedings of the twentieth annual ACM symposium on Principles of distributed computing (PODC '01)* New York: Association for Computing Machinery. *https://doi.org/10.1145/383962.384010*.

98. [HADZILACOS05] Hadzilacos, Vassos. 2005. "On the relationship between the atomic commitment and consensus problems." In *Fault-Tolerant Distributed Computing*, 201-208. London: Springer-Verlag.

99. [HAERDER83] Haerder, Theo, and Andreas Reuter. 1983. "Principles of transaction-oriented database recovery." *ACM Computing Surveys* 15 no. 4 (December):287–317. *https://doi.org/10.1145/289.291*.

100. [HALE10] Hale, Coda. 2010. "You Can't Sacrifice Partition Tolerance." *Coda Hale* (blog). *https://codahale.com/you-cant-sacrifice-partition-tolerance*.

101. [HALPERN90] Halpern, Joseph Y., and Yoram Moses. 1990. "Knowledge and common knowledge in a distributed environment." *Journal of the ACM* 37, no. 3 (July): 549-587. *https://doi.org/10.1145/79147.79161*.

102. [HARDING17] Harding, Rachael, Dana Van Aken, Andrew Pavlo, and Michael Stonebraker. 2017. "An Evaluation of Distributed Concurrency Control." *Proceedings of the VLDB Endowment* 10, no. 5 (January): 553-564. *https://doi.org/10.14778/3055540.3055548*.

103. [HAYASHIBARA04] Hayashibara, N., X. Defago, R.Yared, and T. Katayama. 2004. "The Φ Accrual Failure Detector." In *IEEE Symposium on Reliable Distributed Systems*, 66-78. *https://doi.org/10.1109/RELDIS.2004.1353004*.

104. [HELLAND15] Helland, Pat. 2015. "Immutability Changes Everything." *Queue* 13, no. 9 (November). *https://doi.org/10.1145/2857274.2884038*.

105. [HELLERSTEIN07] Hellerstein, Joseph M., Michael Stonebraker, and James Hamilton. 2007. "Architecture of a Database System." *Foundations and Trends in Databases* 1, no. 2 (February): 141-259. *https://doi.org/10.1561/1900000002*.

106. [HERLIHY94] Herlihy, Maurice. 1994. "Wait-Free Synchronization." *ACM Transactions on Programming Languages and Systems* 13, no. 1 (January): 124-149. *http://dx.doi.org/10.1145/114005.102808*.

107. [HERLIHY10] Herlihy, Maurice, Yossi Lev, Victor Luchangco, and Nir Shavit. 2010. "A Provably Correct Scalable Concurrent Skip List." *https://www.cs.tau.ac.il/~shanir/nir-pubs-web/Papers/OPODIS2006-BA.pdf*.

108. [HERLIHY90] Herlihy, Maurice P., and Jeannette M. Wing. 1990. "Linearizability: a correctness condition for concurrent object." *ACM Transactions on Programming Languages and Systems* 12, no. 3 (July): 463-492. *https://doi.org/10.1145/78969.78972*.

109. [HOWARD14] Howard, Heidi. 2014. "ARC: Analysis of Raft Consensus." Technical Report UCAM-CL-TR-857. Cambridge: University of Cambridge

110. [HOWARD16] Howard, Heidi, Dahlia Malkhi, and Alexander Spiegelman. 2016. "Flexible Paxos: Quorum intersection revisited." *https://arxiv.org/abs/1608.06696*.

111. [HOWARD19] Howard, Heidi, and Richard Mortier. 2019. "A Generalised Solution to Distributed Consensus." *https://arxiv.org/abs/1902.06776*.

112. [HUNT10] Hunt, Patrick, Mahadev Konar, Flavio P. Junqueira, and Benjamin Reed. 2010. "ZooKeeper: wait-free coordination for internet-scale systems." In

Proceedings of the 2010 USENIX conference on USENIX annual technical conference (USENIXATC'10), 11. USENIX.

113. [INTEL14] Intel Corporation. 2014. "Partition Alignment of Intel® SSDs for Achieving Maximum Performance and Endurance." (February). *https://www.intel.com/content/dam/www/public/us/en/documents/technology-briefs/ssd-partition-alignment-tech-brief.pdf*.

114. [JELASITY04] Jelasity, Márk, Rachid Guerraoui, Anne-Marie Kermarrec, and Maarten van Steen. 2004. "The Peer Sampling Service: Experimental Evaluation of Unstructured Gossip-Based Implementations." In *Middleware '04 Proceedings of the 5th ACM/IFIP/USENIX international conference on Middleware*, 79-98. Berlin: Springer-Verlag, Berlin, Heidelberg.

115. [JELASITY07] Jelasity, Márk, Spyros Voulgaris, Rachid Guerraoui, Anne-Marie Kermarrec, and Maarten van Steen. 2007. "Gossip-based Peer Sampling." *ACM Transactions on Computer Systems* 25, no. 3 (August). *http://doi.org/10.1145/1275517.1275520*.

116. [JONSON94] Johnson, Theodore, and Dennis Shasha. 1994. "2Q: A Low Overhead High Performance Buffer Management Replacement Algorithm. " In *VLDB '94 Proceedings of the 20th International Conference on Very Large Data Bases*, 439-450. San Francisco: Morgan Kaufmann.

117. [JUNQUEIRA07] Junqueira, Flavio, Yanhua Mao, and Keith Marzullo. 2007. "Classic Paxos vs. fast Paxos: caveat emptor." In *Proceedings of the 3rd workshop on on Hot Topics in System Dependability (HotDep'07)*. USENIX.

118. [JUNQUEIRA11] Junqueira, Flavio P., Benjamin C. Reed, and Marco Serafini. 2011. "Zab: High-performance broadcast for primary-backup systems." *2011 IEEE/IFIP 41st International Conference on Dependable Systems & Networks (DSN)* (June): 245–256. *https://doi.org/10.1109/DSN.2011.5958223*.

119. [KANNAN18] Kannan, Sudarsun, Nitish Bhat, Ada Gavrilovska, Andrea Arpaci-Dusseau, and Remzi Arpaci-Dusseau. 2018. "Redesigning LSMs for Nonvolatile Memory with NoveLSM." In *USENIX ATC '18 Proceedings of the 2018 USENIX Conference on Usenix Annual Technical Conference*, 993-1005. USENIX.

120. [KARGER97] Karger, D., E. Lehman, T. Leighton, R. Panigrahy, M. Levine, and D. Lewin. 1997. "Consistent hashing and random trees: distributed caching protocols for relieving hot spots on the World Wide Web." In *STOC '97 Proceedings of the twenty-ninth annual ACM symposium on Theory of computing* , 654-663. New York: Association for Computing Machinery.

121. [KEARNEY17] Kearney, Joe. 2017. "Two Phase Commit an old friend." *Joe's Mots* (blog). January 6, 2017. *https://www.joekearney.co.uk/posts/two-phase-commit*.

122. [KEND94] Kendall, Samuel C., Jim Waldo, Ann Wollrath, and Geoff Wyant. 1994. "A Note on Distributed Computing." Technical Report. Mountain View, CA: Sun Microsystems, Inc.

123. [KREMARREC07] Kermarrec, Anne-Marie, and Maarten van Steen. 2007. "Gossiping in distributed systems." *SIGOPS Operating Systems Review* 41, no. 5 (October): 2-7. *https://doi.org/10.1145/1317379.1317381*.

124. [KERRISK10] Kerrisk, Michael. 2010. *The Linux Programming Interface*. San Francisco: No Starch Press.

125. [KHANCHANDANI18] Khanchandani, Pankaj, and Roger Wattenhofer. 2018. "Reducing Compare-and-Swap to Consensus Number One Primitives." *https://arxiv.org/abs/1802.03844*.

126. [KIM12] Kim, Jaehong, Sangwon Seo, Dawoon Jung, Jin-Soo Kim, and Jaehyuk Huh. 2012. "Parameter-Aware I/O Management for Solid State Disks (SSDs)." *IEEE Transactions on Computers* 61, no. 5 (May): 636-649. *https://doi.org/10.1109/TC.2011.76*.

127. [KINGSBURY18a] Kingsbury, Kyle. 2018. "Sequential Consistency." *https://jepsen.io/consistency/models/sequential. 2018.*

128. [KINGSBURY18b] Kingsbury, Kyle. 2018. "Strong consistency models." *Aphyr* (blog). August 8, 2018. *https://aphyr.com/posts/313-strong-consistency-models*.

129. [KLEPPMANN15] Kleppmann, Martin. 2015. "Please stop calling databases CP or AP." *Martin Kleppmann* (blog). May 11, 2015. *https://martin.kleppmann.com/2015/05/11/please-stop-calling-databases-cp-or-ap.html*.

130. [KLEPPMANN14] Kleppmann, Martin, and Alastair R. Beresford. 2014. "A Conflict-Free Replicated JSON Datatype." *https://arxiv.org/abs/1608.03960*.

131. [KNUTH97] Knuth, Donald E. 1997. *The Art of Computer Programming, Volume 1 (3rd Ed.): Fundamental Algorithms*. Boston: Addison-Wesley Longman.

132. [KNUTH98] Knuth, Donald E. 1998. *The Art of Computer Programming, Volume 3: (2nd Ed.): Sorting and Searching*. Boston: Addison-Wesley Longman.

133. [KOOPMAN15] Koopman, Philip, Kevin R. Driscoll, and Brendan Hall. 2015. "Selection of Cyclic Redundancy Code and Checksum Algorithms to Ensure Critical Data Integrity." *U.S. Department of Transportation Federal Aviation Administration*. *https://www.faa.gov/aircraft/air_cert/design_approvals/air_software/media/TC-14–49.pdf*.

134. [KORDAFSHARI05] Kordafshari, M. S., M. Gholipour, M. Mosakhani, A. T. Haghighat, and M. Dehghan. 2005. "Modified bully election algorithm in distributed systems." *Proceedings of the 9th WSEAS International Conference on Computers (ICCOMP'05)*, edited by Nikos E. Mastorakis, Article 10. Stevens Point: World Scientific and Engineering Academy and Society.

135. [KRASKA18] Kraska, Time, Alex Beutel, Ed H. Chi, Jeffrey Dean, and Neoklis Polyzotis. 2018. "The Case for Learned Index Structures." In *SIGMOD '18 Proceedings of the 2018 International Conference on Management of Data*, 489-504. New York: Association for Computing Machinery.

136. [LAMPORT77] Lamport, Leslie. 1977. "Proving the Correctness of Multiprocess Programs." *IEEE Transactions on Software Engineering* 3, no. 2 (March): 125-143. *https://doi.org/10.1109/TSE.1977.229904*.

137. [LAMPORT78] Lamport, Leslie. 1978. "Time, Clocks, and the Ordering of Events in a Distributed System." *Communications of the ACM* 21, no. 7 (July): 558-565

138. [LAMPORT79] Lamport, Leslie. 1979. "How to Make a Multiprocessor Computer That Correctly Executes Multiprocess Programs." *IEEE Transactions on Computers* 28, no. 9 (September): 690-691. *https://doi.org/10.1109/TC.1979.1675439*.

139. [LAMPORT98] Lamport, Leslie. 1998. "The part-time parliament." *ACM Transactions on Computer Systems* 16, no. 2 (May): 133-169. *https://doi.org/10.1145/279227.279229*.

140. [LAMPORT01] Lamport, Leslie. 2001. "Paxos Made Simple." *ACM SIGACT News (Distributed Computing Column)* 32, no. 4 (December): 51-58. *https://www.microsoft.com/en-us/research/publication/paxos-made-simple*.

141. [LAMPORT05] Lamport, Leslie. 2005. "Generalized Consensus and Paxos." *https://www.microsoft.com/en-us/research/publication/generalized-consensus-and-paxos*.

142. [LAMPORT06] Lamport, Leslie. 2006. "Fast Paxos." *Distributed Computing* 19, no. 2 (July): 79-103. *https://doi.org/10.1007/s00446-006-0005-x*.

143. [LAMPORT09] Lamport, Leslie, Dahlia Malkhi, and Lidong Zhou. 2009. "Vertical Paxos and Primary-Backup Replication." In *PODC '09 Proceedings of the 28th ACM symposium on Principles of distributed computing*, 312-313. *https://doi.org/10.1145/1582716.1582783*.

144. [LAMPSON01] Lampson, Butler. 2001. "The ABCD's of Paxos." In *PODC '01 Proceedings of the twentieth annual ACM symposium on Principles of distributed computing*, 13. *https://doi.org/10.1145/383962.383969*.

145. [LAMPSON79] Lampson, Butler W., and Howard E. 1979. "Crash Recovery in a Distributed Data Storage System." *https://www.microsoft.com/en-us/research/publication/crash-recovery-in-a-distributed-data-storage-system'*.

146. [LARRIVEE15] Larrivee, Steve. 2015. "Solid State Drive Primer." *Cactus Technologies* (blog). February 9th, 2015. *https://www.cactus-tech.com/resources/blog/details/solid-state-drive-primer-1-the-basic-nand-flash-cell*.

147. [LARSON81] Larson, Per-Åke, and Åbo Akedemi. 1981. "Analysis of index-sequential files with overflow chaining". *ACM Transactions on Database Systems*. 6, no. 4 (December): 671-680. *https://doi.org/10.1145/319628.319665*.

148. [LEE15] Lee, Collin, Seo Jin Park, Ankita Kejriwal, Satoshi Matsushita, and John Ousterhout. 2015. "Implementing linearizability at large scale and low latency." In *SOSP '15 Proceedings of the 25th Symposium on Operating Systems Principles*, 71-86. *https://doi.org/10.1145/2815400.2815416*.

149. [LEHMAN81] Lehman, Philip L., and s. Bing Yao. 1981. "Efficient locking for concurrent operations on B-trees." *ACM Transactions on Database Systems* 6, no. 4 (December): 650-670. *https://doi.org/10.1145/319628.319663*.

150. [LEITAO07] Leitao, Joao, Jose Pereira, and Luis Rodrigues. 2007. "Epidemic Broadcast Trees." In *SRDS '07 Proceedings of the 26th IEEE International Symposium on Reliable Distributed Systems*, 301-310. IEEE.

151. [LEVANDOSKI14] Levandoski, Justin J., David B. Lomet, and Sudipta Sengupta. 2013. "The Bw-Tree: A B-tree for new hardware platforms." In *Proceedings of the 2013 IEEE International Conference on Data Engineering (ICDE '13)*, 302-313. IEEE. *https://doi.org/10.1109/ICDE.2013.6544834*.

152. [LI10] Li, Yinan, Bingsheng He, Robin Jun Yang, Qiong Luo, and Ke Yi. 2010. "Tree Indexing on Solid State Drives." *Proceedings of the VLDB Endowment* 3, no. 1-2 (September): 1195-1206. *https://doi.org/10.14778/1920841.1920990*.

153. [LIPTON88] Lipton, Richard J., and Jonathan S. Sandberg. 1988. "PRAM: A scalable shared memory." Technical Report, Princeton University. *https://www.cs.princeton.edu/research/techreps/TR-180-88*.

154. [LLOYD11] Lloyd, W., M. J. Freedman, M. Kaminsky, and D. G. Andersen. 2011. "Don't settle for eventual: scalable causal consistency for wide-area storage with COPS." In *Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles (SOSP '11)*, 401-416. New York: Association for Computing Machinery. *https://doi.org/10.1145/2043556.2043593*.

155. [LLOYD13] Lloyd, W., M. J. Freedman, M. Kaminsky, and D. G. Andersen. 2013. "Stronger semantics for low-latency geo-replicated storage." In *10th USENIX Symposium on Networked Systems Design and Implementation (NSDI '13)*, 313-328. USENIX.

156. [LU16] Lu, Lanyue, Thanumalayan Sankaranarayana Pillai, Hariharan Gopalakrishnan, Andrea C. Arpaci-Dusseau, and Remzi H. Arpaci-Dusseau. 2017. "WiscKey: Separating Keys from Values in SSD-Conscious Storage." *ACM Transactions on Storage (TOS)* 13, no. 1 (March): Article 5. *https://doi.org/10.1145/3033273*.

157. [MATTERN88] Mattern, Friedemann. 1988. "Virtual Time and Global States of Distributed Systems." *http://courses.csail.mit.edu/6.852/01/papers/VirtTime_GlobState.pdf*.

158. [MCKENNEY05a] McKenney, Paul E. 2005. "Memory Ordering in Modern Microprocessors, Part I." *Linux Journal* no. 136 (August): 2.

159. [MCKENNEY05b] McKenney, Paul E. 2005. "Memory Ordering in Modern Microprocessors, Part II." *Linux Journal* no. 137 (September): 5.

160. [MEHTA17] Mehta, Apurva, and Jason Gustafson. 2017. "Transactions in Apache Kafka." *Confluent* (blog). November 17, 2017. *https://www.confluent.io/blog/transactions-apache-kafka*.

161. [MELLORCRUMMEY91] Mellor-Crummey, John M., and Michael L. Scott. 1991. "Algorithms for scalable synchronization on shared-memory multiprocessors." *ACM Transactions on Computer Systems* 9, no. 1 (February): 21-65. *https://doi.org/10.1145/103727.103729*.

162. [MELTON06] Melton, Jim. 2006. "Database Language SQL." In *International Organization for Standardization (ISO)*, 105–132. Berlin: Springer. *https://doi.org/10.1007/b137905*.

163. [MERKLE87] Merkle, Ralph C. 1987. "A Digital Signature Based on a Conventional Encryption Function." *A Conference on the Theory and Applications of Cryptographic Techniques on Advances in Cryptology (CRYPTO '87)*, edited by Carl Pomerance. London: Springer-Verlag, 369–378. *https://dl.acm.org/citation.cfm?id=704751*.

164. [MILLER78] Miller, R., and L. Snyder. 1978. "Multiple access to B-trees." *Proceedings of the Conference on Information Sciences and Systems*, Baltimore: Johns Hopkins University (March).

165. [MILOSEVIC11] Milosevic, Z., M. Hutle, and A. Schiper. 2011. "On the Reduction of Atomic Broadcast to Consensus with Byzantine Faults." In *Proceedings of the 2011 IEEE 30th International Symposium on Reliable Distributed Systems (SRDS '11)*, 235-244. IEEE. *https://doi.org/10.1109/SRDS.2011.36*.

166. [MOHAN92] Mohan, C., Don Haderle, Bruce Lindsay, Hamid Pirahesh, and Peter Schwarz. 1992. "ARIES: a transaction recovery method supporting fine-granularity locking and partial rollbacks using write-ahead logging." *Transactions on Database Systems* 17, no. 1 (March): 94-162. *https://doi.org/10.1145/128765.128770*.

167. [MORARU11] Moraru, Iulian, David G. Andersen, and Michael Kaminsky. 2013. "A Proof of Correctness for Egalitarian Paxos." *https://www.pdl.cmu.edu/PDL-FTP/associated/CMU-PDL-13-111.pdf*.

168. [MORARU13] Moraru, Iulian, David G. Andersen, and Michael Kaminsky. 2013. "There Is More Consensus in Egalitarian Parliaments." In *Proceedings of the*

Twenty-Fourth ACM Symposium on Operating Systems Principles (SOSP '13), 358-372. *https://doi.org/10.1145/2517349.2517350*.

169. [MURSHED12] Murshed, Md. Golam, and Alastair R. Allen. 2012. "Enhanced Bully Algorithm for Leader Node Election in Synchronous Distributed Systems." *Computers* 1, no. 1: 3-23. *https://doi.org/10.3390/computers1010003*.

170. [NICHOLS66] Nichols, Ann Eljenholm. 1966. "The Past Participle of 'Overflow:' 'Overflowed' or 'Overflown.'" *American Speech* 41, no. 1 (February): 52–55. *https://doi.org/10.2307/453244*.

171. [NIEVERGELT74] Nievergelt, J. 1974. "Binary search trees and file organization." In *Proceedings of 1972 ACM-SIGFIDET workshop on Data description, access and control (SIGFIDET '72)*, 165-187. *https://doi.org/10.1145/800295.811490*.

172. [NORVIG01] Norvig, Peter. 2001. "Teach Yourself Programming in Ten Years." *https://norvig.com/21-days.html*.

173. [ONEIL93] O'Neil, Elizabeth J., Patrick E. O'Neil, and Gerhard Weikum. 1993. "The LRU-K page replacement algorithm for database disk buffering." In *Proceedings of the 1993 ACM SIGMOD international conference on Management of data (SIGMOD '93)*, 297-306. *https://doi.org/10.1145/170035.170081*.

174. [ONEIL96] O'Neil, Patrick, Edward Cheng, Dieter Gawlick, and Elizabeth O'Neil. 1996. "The log-structured merge-tree (LSM-tree)." *Acta Informatica* 33, no. 4: 351-385. *https://doi.org/10.1007/s002360050048*.

175. [ONGARO14] Ongaro, Diego and John Ousterhout. 2014. "In Search of an Understandable Consensus Algorithm." In *Proceedings of the 2014 USENIX conference on USENIX Annual Technical Conference (USENIX ATC'14)*, 305-320. USENIX.

176. [OUYANG14] Ouyang, Jian, Shiding Lin, Song Jiang, Zhenyu Hou, Yong Wang, and Yuanzheng Wang. 2014. "SDF: software-defined flash for web-scale internet storage systems." *ACM SIGARCH Computer Architecture News* 42, no. 1 (February): 471-484. *https://doi.org/10.1145/2654822.2541959*.

177. [PAPADAKIS93] Papadakis, Thomas. 1993. "Skip lists and probabilistic analysis of algorithms." Doctoral Dissertation, University of Waterloo. *https://cs.uwaterloo.ca/research/tr/1993/28/root2side.pdf*.

178. [PUGH90a] Pugh, William. 1990. "Concurrent Maintenance of Skip Lists." Technical Report, University of Maryland. *https://drum.lib.umd.edu/handle/1903/542*.

179. [PUGH90b] Pugh, William. 1990. "Skip lists: a probabilistic alternative to balanced trees." *Communications of the ACM* 33, no. 6 (June): 668-676. *https://doi.org/10.1145/78973.78977*.

180. [RAMAKRISHNAN03] Ramakrishnan, Raghu, and Johannes Gehrke. 2002. *Database Management Systems (3rd Ed.)*. New York: McGraw-Hill.

181. [RAY95] Ray, Gautam, Jayant Haritsa, and S. Seshadri. 1995. "Database Compression: A Performance Enhancement Tool." In *Proceedings of 7th International Conference on Management of Data (COMAD)*. New York: McGraw Hill.

182. [RAYNAL99] Raynal, M., and F. Tronel. 1999. "Group membership failure detection: a simple protocol and its probabilistic analysis." *Distributed Systems Engineering* 6, no. 3 (September): 95-102. *https://doi.org/10.1088/0967-1846/6/3/301*.

183. [REED78] Reed, D. P. 1978. "Naming and synchronization in a decentralized computer system." Technical Report, MIT. *https://dspace.mit.edu/handle/1721.1/16279*.

184. [REN16] Ren, Kun, Jose M. Faleiro, and Daniel J. Abadi. 2016. "Design Principles for Scaling Multi-core OLTP Under High Contention." In *Proceedings of the 2016 International Conference on Management of Data (SIGMOD '16)*, 1583-1598. *https://doi.org/10.1145/2882903.2882958*.

185. [ROBINSON08] Robinson, Henry. 2008. "Consensus Protocols: Two-Phase Commit." *The Paper Trail* (blog). November 27, 2008. *https://www.the-paper-trail.org/post/2008-11-27-consensus-protocols-two-phase-commit*.

186. [ROSENBLUM92] Rosenblum, Mendel, and John K. Ousterhout. 1992. "The Design and Implementation of a Log Structured File System." *ACM Transactions on Computer Systems* 10, no. 1 (February): 26-52. *https://doi.org/10.1145/146941.146943*.

187. [ROY12] Roy, Arjun G., Mohammad K. Hossain, Arijit Chatterjee, and William Perrizo. 2012. "Column-oriented Database Systems: A Comparison Study." In *Proceedings of the ISCA 27th International Conference on Computers and Their Applications*, 264-269.

188. [RUSSEL12] Russell, Sears. 2012. "A concurrent skiplist with hazard pointers." *http://rsea.rs/skiplist*.

189. [RYSTSOV16] Rystsov, Denis. 2016. "Best of both worlds: Raft's joint consensus + Single Decree Paxos." *Rystsov.info* (blog). January 5, 2016. *http://rystsov.info/2016/01/05/raft-paxos.html*.

190. [RYSTSOV18] Rystsov, Denis. 2018. "Replicated State Machines without logs." *https://arxiv.org/abs/1802.07000*.

191. [SATZGER07] Satzger, Benjamin, Andreas Pietzowski, Wolfgang Trumler, and Theo Ungerer. 2007. "A new adaptive accrual failure detector for dependable distributed systems." In *Proceedings of the 2007 ACM symposium on Applied computing (SAC '07)*, 551-555. *https://doi.org/10.1145/1244002.1244129*.

192. [SAVARD05] Savard, John. 2005. "Floating-Point Formats." *http://www.quadibloc.com/comp/cp0201.htm*.

193. [SCHWARZ86] Schwarz, P., W. Chang, J. C. Freytag, G. Lohman, J. McPherson, C. Mohan, and H. Pirahesh. 1986. "Extensibility in the Starburst database system." In *OODS '86 Proceedings on the 1986 international workshop on Object-oriented database systems*, 85–92. IEEE.

194. [SEDGEWICK11] Sedgewick, Robert, and Kevin Wayne. 2011. *Algorithms (4th Ed.)*. Boston: Pearson.

195. [SHAPIRO11a] Shapiro, Marc, Nuno Preguiça, Carlos Baquero, and Marek Zawirski. 2011. "Conflict-free Replicated Data Types." In *Stabilization, Safety, and Security of Distributed Systems*, 386-400. Berlin: Springer, Berlin, Heidelberg.

196. [SHAPIRO11b] Shapiro, Marc, Nuno Preguiça, Carlos Baquero, and Marek Zawirski. 2011. "A comprehensive study of Convergent and Commutative Replicated Data Types." *https://hal.inria.fr/inria-00555588/document*.

197. [SHEEHY10a] Sheehy, Justin. 2010. "Why Vector Clocks Are Hard." *Riak* (blog). April 5, 2010. *https://riak.com/posts/technical/why-vector-clocks-are-hard*.

198. [SHEEHY10b] Sheehy, Justin, and David Smith. 2010. "Bitcask, A Log-Structured Hash Table for Fast Key/Value Data."

199. [SILBERSCHATZ10] Silberschatz, Abraham, Henry F. Korth, and S. Sudarshan. 2010. *Database Systems Concepts (6th Ed.)*. New York: McGraw-Hill.

200. [SINHA97] Sinha, Pradeep K. 1997. *Distributed Operating Systems: Concepts and Design*. Hoboken, NJ: Wiley.

201. [SKEEN82] Skeen, Dale. 1982. "A Quorum-Based Commit Protocol." Technical Report, Cornell University.

202. [SKEEN83] Skeen, Dale, and M. Stonebraker. 1983. "A Formal Model of Crash Recovery in a Distributed System." *IEEE Transactions on Software Engineering* 9, no. 3 (May): 219-228. *https://doi.org/10.1109/TSE.1983.236608*.

203. [SOUNDARARARJAN06] Soundararajan, Gokul. 2006. "Implementing a Better Cache Replacement Algorithm in Apache Derby Progress Report." *https://pdfs.semanticscholar.org/220b/2fe62f13478f1ec75cf17ad085874689c604.pdf*.

204. [STONE98] Stone, J., M. Greenwald, C. Partridge and J. Hughes. 1998. "Performance of checksums and CRCs over real data." *IEEE/ACM Transactions on Networking* 6, no. 5 (October): 529-543. *https://doi.org/10.1109/90.731187*.

205. [TANENBAUM14] Tanenbaum, Andrew S., and Herbert Bos. 2014. *Modern Operating Systems (4th Ed.)*. Upper Saddle River: Prentice Hall Press.

206. [TANENBAUM06] Tanenbaum, Andrew S., and Maarten van Steen. 2006. *Distributed Systems: Principles and Paradigms*. Boston: Pearson.

207. [TARIQ11] Tariq, Ovais. 2011. "Understanding InnoDB clustered indexes." *Ovais Tariq* (blog). January 20, 2011. *http://www.ovaistariq.net/521/understanding-innodb-clustered-indexes/#.XTtaUpNKj5Y*.

208. [TERRY94] Terry, Douglas B., Alan J. Demers, Karin Petersen, Mike J. Spreitzer, Marvin M. Theimer, and Brent B. Welch. 1994. "Session Guarantees for Weakly Consistent Replicated Data." In *PDIS '94 Proceedings of the Third International Conference on Parallel and Distributed Information Systems*, 140–149. IEEE.

209. [THOMAS79] Thomas, Robert H. 1979. "A majority consensus approach to concurrency control for multiple copy databases." *ACM Transactions on Database Systems* 4, no. 2 (June): 180–209. *https://doi.org/10.1145/320071.320076*.

210. [THOMSON12] Thomson, Alexander, Thaddeus Diamond, Shu-Chun Weng, Kun Ren, Philip Shao, and Daniel J. Abadi. 2012. "Calvin: Fast distributed transactions for partitioned database systems." In *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '12)*. New York: Association for Computing Machinery. *https://doi.org/10.1145/2213836.2213838*.

211. [VANRENESSE98] van Renesse, Robbert, Yaron Minsky, and Mark Hayden. 1998. "A Gossip-Style Failure Detection Service." In *Middleware '98 Proceedings of the IFIP International Conference on Distributed Systems Platforms and Open Distributed Processing*, 55–70. London: Springer-Verlag.

212. [VENKATARAMAN11] Venkataraman, Shivaram, Niraj Tolia, Parthasarathy Ranganathan, and Roy H. Campbell. 2011. "Consistent and Durable Data Structures for Non-Volatile Byte-Addressable Memory." In *Proceedings of the 9th USENIX conference on File and stroage technologies (FAST'11)*, 5. USENIX.

213. [VINOSKI08] Vinoski, Steve. 2008. "Convenience Over Correctness." *IEEE Internet Computing* 12, no. 4 (August): 89–92. *https://doi.org/10.1109/MIC.2008.75*.

214. [VIOTTI16] Viotti, Paolo, and Marko Vukolić. 2016. "Consistency in Non-Transactional Distributed Storage Systems." *ACM Computing Surveys* 49, no. 1 (July): Article 19. *https://doi.org/0.1145/2926965*.

215. [VOGELS09] Vogels, Werner. 2009. "Eventually consistent." *Communications of the ACM* 52, no. 1 (January): 40–44. *https://doi.org/10.1145/1435417.1435432*.

216. [WALDO96] Waldo, Jim, Geoff Wyant, Ann Wollrath, and Samuel C. Kendall. 1996. "A Note on Distributed Computing." *Selected Presentations and Invited Papers SecondInternational Workshop on Mobile Object Systems—Towards the Programmable Internet* (July): 49–64. *https://dl.acm.org/citation.cfm?id=747342*.

217. [WANG13] Wang, Peng, Guangyu Sun, Song Jiang, Jian Ouyang, Shiding Lin, Chen Zhang, and Jason Cong. 2014. "An Efficient Design and Implementation of LSM-tree based Key-Value Store on Open-Channel SSD." *EuroSys '14 Proceedings of the Ninth European Conference on Computer Systems* (April): Article 16. *https://doi.org/10.1145/2592798.2592804*.

218. [WANG18] Wang, Ziqi, Andrew Pavlo, Hyeontaek Lim, Viktor Leis, Huanchen Zhang, Michael Kaminsky, and David G. Andersen. 2018. "Building a Bw-Tree Takes More Than Just Buzz Words." *Proceedings of the 2018 International Confer-*

ence on Management of Data (SIGMOD '18), 473–488. *https://doi.org/10.1145/3183713.3196895*.

219. [WEIKUM01] Weikum, Gerhard, and Gottfried Vossen. 2001. *Transactional Information Systems: Theory, Algorithms, and the Practice of Concurrency Control and Recovery*. San Francisco: Morgan Kaufmann Publishers Inc.

220. [XIA17] Xia, Fei, Dejun Jiang, Jin Xiong, and Ninghui Sun. 2017. "HiKV: A Hybrid Index Key-Value Store for DRAM-NVM Memory Systems." *Proceedings of the 2017 USENIX Annual Technical Conference (USENIX ATC '17)*, 349–362. USENIX.

221. [YANG14] Yang, Jingpei, Ned Plasson, Greg Gillis, Nisha Talagala, and Swaminathan Sundararaman. 2014. "Don't stack your Log on my Log." *INFLOW* (October). *https://www.usenix.org/system/files/conference/inflow14/inflow14-yang.pdf*.

222. [ZHAO15] Zhao, Wenbing. 2015. "Fast Paxos Made Easy: Theory and Implementation." *International Journal of Distributed Systems and Technologies* 6, no. 1 (January): 15-33. *https://doi.org/10.4018/ijdst.2015010102*.