

















































- [85] Baker J, Bond C, Corbett JC, Furman J, Khorlin A, Larson J, Leon JM, Li Y, Lloyd A, Yushprakh V. Megastore: Providing scalable, highly available storage for interactive services. In: Proc. of the Conf. on Innovative Data Systems Research (CIDR 2013). 2011. 223–234. [http://cidrdb.org/cidr2011/Papers/CIDR11\\_Paper32.pdf](http://cidrdb.org/cidr2011/Papers/CIDR11_Paper32.pdf)
- [86] Peng D, Dabek F. Large-Scale incremental processing using distributed transactions and notifications. In: Proc. of the USENIX Conf. on Operating Systems Design and Implementation (OSDI 2010). Vancouver: USENIX Association, 2010. 1–15. [https://www.usenix.org/legacy/event/osdi10/tech/full\\_papers/Peng.pdf](https://www.usenix.org/legacy/event/osdi10/tech/full_papers/Peng.pdf)
- [87] Krueger J, Kim C, Grund M, Satish N, Schwalb D, Chhugani J, Plattner H, Dubey P, Zeier A. Fast updates on read-optimized databases using multi-core CPUs. Proc. of the VLDB Endowment, 2011,5(1):61–72. [doi: 10.14778/2047485.2047491]
- [88] VoltDB Team. VoltDB. 2016. <https://www.voltdb.com/>
- [89] Corbett JC, Dean J, Epstein M, Fikes A, Frost C, Furman JJ, Ghemawat S, Gubarev A, Heiser C, Hochschild P, Hsieh W, Kanthak S, Kogan E, Li H, Lloyd A, Melnik S, Mwaura D, Nagle D, Quinlan S, Rao R, Rolig L, Saito Y, Szymaniak M, Taylor C, Wang R, Woodford D. Spanner: Google’s globally-distributed database. In: Proc. of the 10th USENIX Conf. on Operating Systems Design and Implementation (OSDI 2012). Hollywood: USENIX Association, 2012. 251–264. <https://www.usenix.org/system/files/conference/osdi12/osdi12-final-16.pdf>
- [90] CockroachDB Team. CockroachDB. 2016. <https://www.cockroachlabs.com/>
- [91] Yuan LY, Wu L, You JH, Chi Y. A demonstration of rubato DB: A highly scalable newSQL database system for OLTP and big data applications. In: Proc. of the 2015 ACM SIGMOD Int’l Conf. on Management of Data (SIGMOD 2015). Melbourne: ACM Press, 2015. 907–912. [doi: 10.1145/2723372.2735380]
- [92] Wu L, Yuan LY, You JH. BASIC: An alternative to BASE for large-scale data management system. In: Proc. of the IEEE Int’l Conf. on Big Data (IEEE Big Data 2014). 2014. 5–14. [doi: 10.1109/BigData.2014.7004206]
- [93] DeBrabant J, Pavlo A, Tu S, Stonebraker M, Zdonik S. Anti-Caching: A new approach to database management system architecture. Proc. of the VLDB Endowment (VLDB 2013), 2013,6(14):1942–1953. [doi: 10.14778/2556549.2556575]
- [94] Stoica R, Ailamaki A. Enabling efficient OS paging for main-memory OLTP databases. In: Proc. of the 9th Int’l Workshop on Data Management on New Hardware (DaMoN 2013). New York: ACM Press, 2013. 1–7. [doi: 10.1145/2485278.2485285]
- [95] Pavlo A. Emerging hardware trends in large-scale transaction processing. IEEE Internet Computing, 2015,19(3):68–71. [doi: 10.1109/MIC.2015.59]
- [96] Stonebraker M. Errors in database systems, eventual consistency, and the CAP theorem. 2010. <http://cacm.acm.org/blogs/blog-cacm/83396-errors-in-database-systems-eventual-consistency-and-the-cap-theorem/fulltext>.
- [97] Abadi D. Consistency tradeoffs in modern distributed database system design: CAP is only part of the story. Computer, 2012,45(2): 37–42. [doi: 10.1109/mc.2012.33]
- [98] Zhenkun Y, Chuanhui Y, Zhen L. OceanBase—A massive structured data storage management system. E-science Technology & Application, 2013,4(1):41–48.
- [99] Bailis P, Davidson A, Fekete A, Ghodsi A, Hellerstein JM, Stoica I. Highly available transactions: Virtues and limitations. Proc. of the VLDB Endowment, 2013,7(3):181–192. [doi: 10.14778/2732232.2732237]
- [100] Campos AF, Esteves S, Veiga L. HBase++: Extending HBase with client-centric consistency guarantees for geo-replication. 2013. <http://www.gsd.inesc-id.pt/~sesteves/papers/inforum14-hbase-plus-plus.pdf>
- [101] Helland P. Life beyond distributed transactions: An apostate’s opinion. In: Proc. of the Biennial Conf. on Innovative Data Systems Research (CIDR 2007). Asilomar, 2007. 132–141. <http://cidrdb.org/cidr2007/papers/cidr07p15.pdf>
- [102] Stonebraker M, Madden S, Abadi DJ, Harizopoulos S, Hachem N, Helland P. The end of an architectural era (it’s time for a complete rewrite). In: Proc. of the 33rd Int’l Conf. on Very Large Data Bases (VLDB 2007). Vienna: VLDB Endowment, 2007. 1150–1160. <http://www.cs.yale.edu/homes/dna/vldb07hstore.pdf>
- [103] Pavlo A, Curino C, Zdonik S. Skew-Aware automatic database partitioning in shared-nothing, parallel OLTP systems. In: Proc. of the 2012 ACM SIGMOD Int’l Conf. on Management of Data (SIGMOD 2012). Scottsdale: ACM Press, 2012. 61–72. [doi: 10.1145/2213836.2213844]
- [104] NuoDB Team. NuoDB. 2016. <http://www.nuodb.com/>
- [105] MemSQL Team. MemSQL. 2016. <http://www.memsql.com/>

- [106] Kemme B, Alonso G. Don't be lazy, be consistent: Postgres-R, a new way to implement database replication. In: Proc. of the 26th Int'l Conf. on Very Large Data Bases (VLDB 2000). Morgan Kaufmann Publishers, 2000. 134–143. [https://static.aminer.org/pdf/PDF/000/642/954/don\\_t\\_be\\_lazy\\_be\\_consistent\\_postgres\\_r\\_a\\_new.pdf](https://static.aminer.org/pdf/PDF/000/642/954/don_t_be_lazy_be_consistent_postgres_r_a_new.pdf)
- [107] Jiminez-Peris R, Patino-Martinez M, Arevalo S. Deterministic scheduling for transactional multithreaded replicas. In: Proc. of the 19th IEEE Symp. on Reliable Distributed Systems (SRDS 2000). 2000. 164–173. [doi: 10.1109/RELDI.2000.885404]
- [108] Thomson A, Diamond T, Weng SC, Ren K, Shao P, Abadi DJ. Calvin: Fast distributed transactions for partitioned database systems. In: Proc. of the 2012 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD 2012). Scottsdale: ACM Press, 2012. 1–12. [doi: 10.1145/2213836.2213838]
- [109] Ren K, Thomson A, Abadi DJ. An evaluation of the advantages and disadvantages of deterministic database systems. Proc. of the VLDB Endowment, 2014,7(10):821–832. [doi: 10.14778/2732951.2732955]
- [110] Sikka V, Farber F, Goel A, Lehner W. SAP HANA: The evolution from a modern main-memory data platform to an enterprise application platform. Proc. of the VLDB Endowment, 2013,6(11):1184–1185. [doi: 10.14778/2536222.2536251]
- [111] Ferro DG, Junqueira F, Kelly I, Reed B, Yabandeh M. Omid: Lock-Free transactional support for distributed data stores. In: Proc. of the 2014 IEEE 30th Int'l Conf. on Data Engineering (ICDE 2014). 2014. 676–687. [doi: 10.1109/ICDE.2014.6816691]
- [112] Elmore AJ, Arora V, Taft R, Pavlo A, Agrawal D, Abbadi AE. Squall: Fine-Grained live reconfiguration for partitioned main memory databases. In: Proc. of the 2015 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD 2015). Melbourne: ACM Press, 2015. 299–313. [doi: 10.1145/2723372.2723726]
- [113] DeBrabant J, Arulraj J, Pavlo A, Stonebraker M, Zdonik SB, Dullloor S. A prolegomenon on OLTP database systems for non-volatile memory. In: Proc. of the 5th Int'l Workshop on Accelerating Data Management Systems Using Modern Processor and Storage Architectures (ADMS). Hangzhou, 2014. 57–63. [http://adms-conf.org/2014/adms14\\_debrabant.pdf](http://adms-conf.org/2014/adms14_debrabant.pdf)
- [114] Coburn J, Bunker T, Schwarz M, Gupta R, Swanson S. From ARIES to MARS: Transaction support for next-generation, solid-state drives. In: Proc. of the 24th ACM Symp. on Operating Systems Principles (SOSP 2013). Farminton: ACM Press, 2013. 197–212. [doi: 10.1145/2517349.2522724]
- [115] Pelley S, Wenisch TF, Gold BT, Bridge B. Storage management in the NVRAM era. Proc. of the VLDB Endowment, 2013,7(2): 121–132. [doi: 10.14778/2732228.2732231]
- [116] Gao S, Xu J, Haerder T, He B, Choi B, Hu H. PCMLogging: Optimizing transaction logging and recovery performance with PCM. IEEE Trans. on Knowledge & Data Engineering, 2015,27(12):3332–3346. [doi: 10.1109/TKDE.2015.2453154]
- [117] Arulraj J, Pavlo A, Dullloor SR. Let's talk about storage & recovery methods for non-volatile memory database systems. In: Proc. of the 2015 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD 2015). Melbourne: ACM Press, 2015. 707–722. [doi: 10.1145/2723372.2749441]
- [118] Huang J, Schwan K, Qureshi MK. NVRAM-Aware logging in transaction systems. Proc. of the VLDB Endowment, 2014,8(4): 389–400. [doi: 10.14778/2735496.2735502]
- [119] Wang T, Johnson R. Scalable logging through emerging non-volatile memory. Proc. of the VLDB Endowment, 2014,7(10): 865–876. [doi: 10.14778/2732951.2732960]
- [120] Oukid I, Booss D, Lehner W, Bumbulis P, Willhalm T. SOFORT: A hybrid SCM-DRAM storage engine for fast data recovery. In: Proc. of the 10th Int'l Workshop on Data Management on New Hardware (DaMoN 2014). Snowbird: ACM Press, 2014. 1–7. [doi: 10.1145/2619228.2619236]
- [121] Kolli A, Pelley S, Saidi A, Chen PM, Wenisch TF. High-Performance transactions for persistent memories. SIGARCH Computer Architecture News, 2016,44(2):399–411. [doi: 10.1145/2980024.2872381]
- [122] Larson PK, Blanas S, Diaconu C, Freedman C, Patel JM, Zwilling M. High-Performance concurrency control mechanisms for main-memory databases. Proc. of the VLDB Endowment, 2011,5(4):298–309. [doi: 10.14778/2095686.2095689]
- [123] Johnson R, Pandis I, Ailamaki A. Improving OLTP scalability using speculative lock inheritance. Proc. of the VLDB Endowment, 2009,2(1):479–489. [doi: 10.14778/1687627.1687682]
- [124] Pandis I, Johnson R, Hardavellas N, Ailamaki A. Data-Oriented transaction execution. Proc. of the VLDB Endowment, 2010,3(1-2): 928–939. [doi: 10.14778/1920841.1920959]



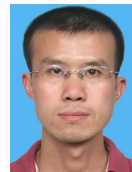
- [125] Atkins MS, Coady MY. Adaptable concurrency control for atomic data types. *ACM Trans. on Computer Systems*, 1992,10(3): 190–225. [doi: 10.1145/146937.146939]
- [126] Joshi AM. Adaptive locking strategies in a multi-node data sharing environment. In: *Proc. of the 17th Int'l Conf. on Very Large Data Bases (VLDB'91)*. Morgan Kaufmann Publishers, 1991. 181–191. <http://vldb.org/conf/1991/P181.PDF>
- [127] Ren K, Thomson A, Abadi DJ. Lightweight locking for main memory database systems. *Proc. of the VLDB Endowment*, 2012,6(2): 145–156. [doi: 10.14778/2535568.2448947]
- [128] Ren K, Thomson A, Abadi DJ. VLL: A lock manager redesign for main memory database systems. *The VLDB Journal*, 2015,24(5): 681–705. [doi: 10.1007/s00778-014-0377-7]
- [129] Sadoghi M, Ross KA, Canim M, Bhattacharjee B. Making updates disk-I/O friendly using SSDs. *Proc. of the VLDB Endowment*, 2013,6(11):997–1008. [doi: 10.14778/2536222.2536226]
- [130] Ailamaki A. Databases and hardware: The beginning and sequel of a beautiful friendship. *Proc. of the VLDB Endowment*, 2015, 8(12):2058–2061. [doi: 10.14778/2824032.2824142]
- [131] Viglas SD. Data management in non-volatile memory. In: *Proc. of the 2015 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD 2015)*. Melbourne: ACM Press, 2015. 1707–1711. [doi: 10.1145/2723372.2731082]
- [132] Viglas SD. Write-Limited sorts and joins for persistent memory. *Proc. of the VLDB Endowment*, 2014,7(5):413–424. [doi: 10.14778/2732269.2732277]
- [133] Chen S, Gibbons PB, Nath S. Rethinking database algorithms for phase change memory. In: *Proc. of the Conf. on Innovative Data Systeems Research (CIDR 2011)*. 2011. 21–31. [http://cidrdb.org/cidr2011/Papers/CIDR11\\_Paper3.pdf](http://cidrdb.org/cidr2011/Papers/CIDR11_Paper3.pdf)
- [134] Ren K, Faleiro JM, Abadi DJ. Design principles for scaling multi-core OLTP under high contention. In: *Proc. of the 2016 Int'l Conf. on Management of Data (SIGMOD 2016)*. San Francisco: ACM Press, 2016. 1583–1598. [doi: 10.1145/2882903.2882958]
- [135] Zhu YA, Zhou X, Zjang YS. A survey of optimization methods for transactional database in multi-core era. *Chinese Journal of Computers*, 2015,38(9):1865–1879 (in Chinese with English abstract). [doi: 10.11897/SP.J.1016.2015.01865]

#### 附中文参考文献:

- [2] 罗乐,刘轶,钱德沛.内存计算技术研究综述. *软件学报*,2016,27(8):2147–2167. <http://www.jos.org.cn/1000-9825/5103.htm> [doi: 10.13328/j.cnki.jos.005103]
- [135] 朱阅岸,周烜,张延松,周明,牛嘉,王珊.多核处理器下事务型数据库性能优化技术综述. *计算机学报*,2015,38(9):1865–1879. [doi: 10.11897/SP.J.1016.2015.01865]



潘巍(1977—),男,安徽蚌埠人,博士,副教授,CCF 专业会员,主要研究领域为数据库理论与技术,海量数据管理,内存计算。



周陈超(1982—),男,博士,工程师,CCF 学生会员,主要研究领域为知识发现,知识管理。



李战怀(1961—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为数据库理论与技术,海量数据存储与管理。



苏静(1987—),女,博士,主要研究领域为大数据处理技术,图数据处理。



杜洪涛(1978—),男,博士,讲师,主要研究领域为海量数据管理,分布式数据库。