The Missing Dimensions in Geo-Distributed Database Evaluation

Oto Mraz [1], Kyriakos Psarakis [1], George Christodoulou [1], Paris Carbone [2], Asterios Katsifodimos [1]

[1] Software Technology Department, Delft University of Technology, Netherlands
{o.m.mraz,k.psarakis,g.c.christodoulou,a.katsifodimos}@tudelft.nl
[2] Department of Computer Science, KTH Royal Institute of Technology, Sweden
parisc@kth.se

Geo-distributed transactional databases underpin modern applications in banking, e-commerce, and financial trading. By partitioning and replicating data across different geographical regions, they enhance availability, tolerate regional failures, and reduce user-perceived latency. However, ensuring strong consistency and atomicity across regions requires complex coordination protocols that may involve multiple rounds of communication. Transactions may access data locally or span multiple regions, and can even be submitted from regions that do not store the data they access, introducing complex performance trade-offs that recent systems seek to optimize [4, 6, 5, 3].

However, current evaluation practices often fail to capture the realities of wide-area deployments. They typically assume uniform and stable networks, and overlook data and client locality [6, 5, 3, 2]. In practice, transactions may incur significant overhead compared to single-region deployments due to cross-region communication latency across a wide-area network (WAN). Besides impeding database performance, WAN communication is also very costly in public cloud deployments. Despite this, data transfer volumes and their monetary implications are rarely reported. Likewise, evaluations seldom compare performance across hardware types, or benchmark on workloads reflecting a range of key geo-distributed access patterns such as multi-home, local-single-home, and foreign-single-home transactions [1, 7].

In this work, we revisit existing benchmarking practices of geo-distributed OLTP systems and expose key gaps in modeling real-world behavior. To address these shortcomings, we developed Gaia, a benchmarking framework that enables a comprehensive evaluation of state-of-the-art databases across eight cloud regions. Gaia introduces a set of new evaluation dimensions: transaction locality, submission-placement asymmetry, hardware heterogeneity, and cross-region network variability. Our framework also augments traditional performance metrics (throughput and latency) with data transfers and monetary cost per transaction. Our experiments reveal that *i*) cross-region transfers dominate the deployment cost in the cloud, *ii*) all systems are highly sensitive to network instability, and *iii*) optimal hardware selection is a non-trivial trade-off between throughput, latency, and monetary cost. We argue that for the design of a reliable and efficient geo-distributed database, we must rethink how performance and cost can be optimized simultaneously.

References

[1] R. Harding et al. "An evaluation of distributed concurrency control". In: *Proc. VLDB Endowment* 10.5 (2017), pp. 553–564.

[2] J. Hildred, M. Abebe, and K. Daudjee. "Caerus: Low-Latency Distributed Transactions for Geo-Replicated Systems". In: *Proceedings of the VLDB Endowment* 17.3 (2023), pp. 469–482.

[3] S. Mu et al. "Consolidating concurrency control and consensus for commits under conflicts". In: *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*. 2016, pp. 517–532.

[4] C. D. Nguyen, J. K. Miller, and D. J. Abadi. "Detock: High performance multi-region transactions at scale". In: *Proc. ACM on Management of Data* 1.2 (2023), pp. 1–27.

[5] K. Ren, D. Li, and D. J. Abadi. "SLOG: Serializable, low-latency, geo-replicated transactions". In: *Proc. VLDB Endowment* 12.11 (2019).

[6] A. Thomson et al. "Calvin: Fast distributed transactions for partitioned database systems". In: *Proc. 2012 ACM SIGMOD International Conference on Management of Data*. 2012, pp. 1–12.

[7] X. Yu et al. "Staring into the Abyss: An Evaluation of Concurrency Control with One Thousand Cores". In: *Proceedings of the VLDB Endowment* 8.3 (2014).