DeathStar Movie for Geo-Distributed Databases

S. E. van den Houten

Supervisors: Oto Mráz, Dr. Asterios Katsifodimos

Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology

1 Background

Geo-distributed databases power many critical systems. It is therefore important that they are tested well. However, industry-standard benchmarks for these systems are currently not sufficient[2]. In this research, we explore the potential of using the DeathStar Movie benchmark[1] for testing geo-distributed databases.

2 Research question

- **How do geo-distributed databases perform on the DeathStar Movie benchmark?**

3 DeathStar Movie for databases

DeathStar Movie was originally designed to test microservice systems. We modified it to extract just the relevant database transactions, resulting in the following database schema:

users	
username 🔑	varchar
first_name	varchar
last_name	varchar
password	varchar
user_id	integer
reviews	integer

reviews	
review_id 🔑	integer
user_id	integer
movie_id	varchar
req_id	integer
text	varchar
rating	integer
timestamp	integer

| movies	
title 🔑	varchar
movie_id	varchar

The workload consists of the *NewReview* transaction, which fetches user and movie data and then creates a review.
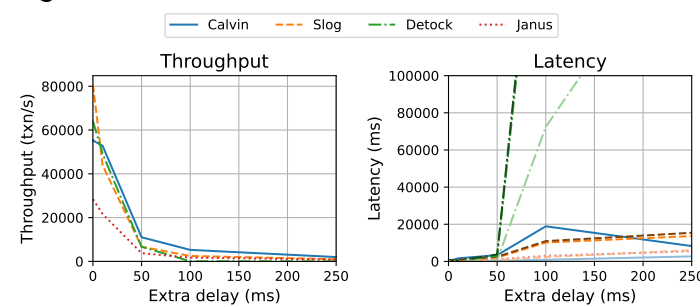
4 Experiment

The experiment consisted of running six variants of the benchmark on four different geo-distributed databases.
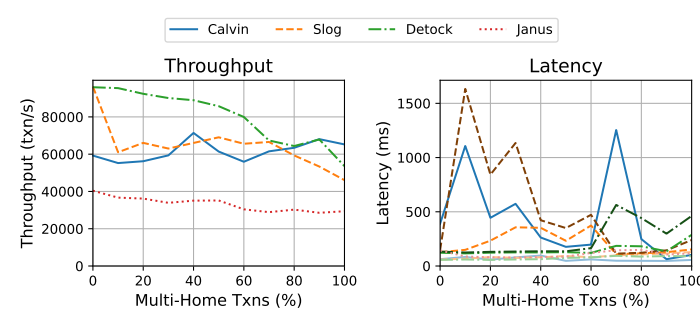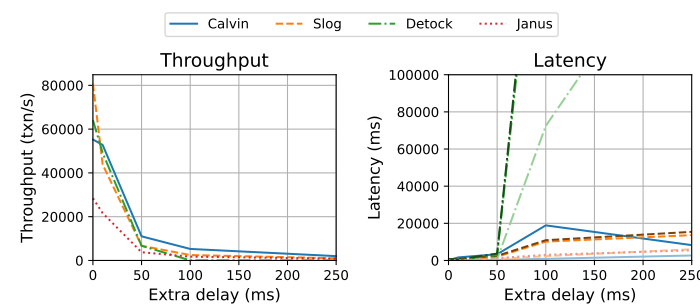

The **Baseline** scenario shows that SLOG and Detock can be outperformed by the "primitive" Calvin when hit with many multi-region transactions.
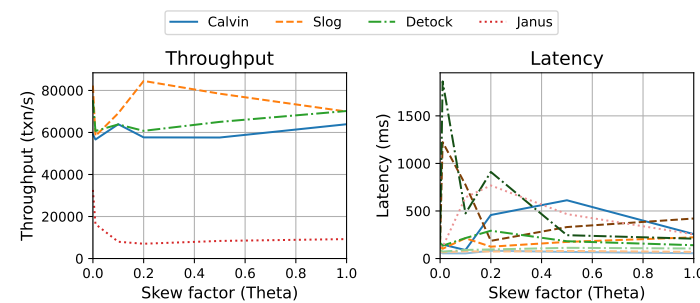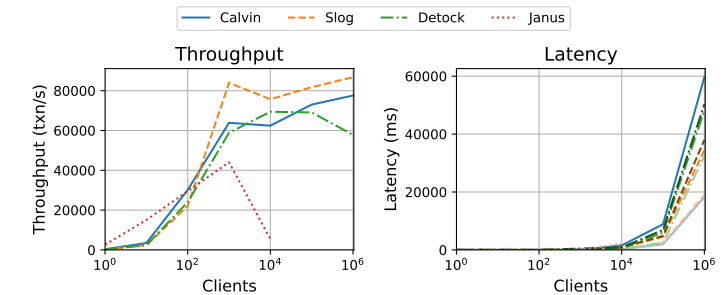

The **Skew** scenario shows the particular weakness of Janus to high contention
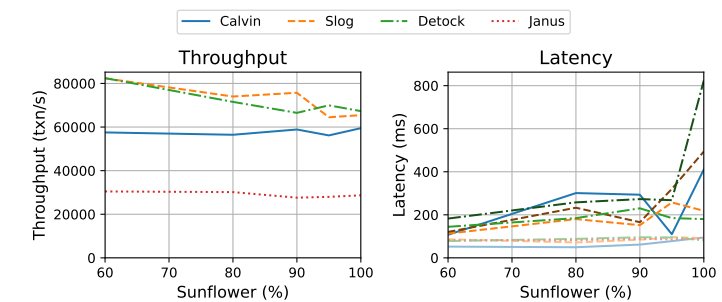

The **Network Latency** scenario shows Detock's dramatic performance reduction under increased latency.

The **Packet Loss** scenario also sees Detock suffer when limited by the network.

The **Scalability** scenario reveals a performance sweet spot at around 10^3 clients.

The **Sunflower** scenario shows performance degradation when data usage is skewed to one region.

5 Future work

- Experiments combining the effects of multiple scenarios.
- Experiments on a more realistic testing platform.
- Investigate Detock performance issues when running on a limited network.

6 Conclusion

DeathStar Movie can be a valuable tool for benchmarking geo-distributed databases. The experiments showed the strengths and weaknesses of the different database systems. Particularly, it showed Janus' struggles with high contention and Detock's reliance on a performant network.

References

[1] Y. Gan, Y. Zhang, D. Cheng, A. Shetty, P. Rathi, N. Katarki, A. Bruno, J. Hu, B. Ritchken, B. Jackson, K. Hu, M. Pancholi, Y. He, B. Clancy, C. Colen, F. Wen, C. Leung, S. Wang, L. Zaruvinsky, M. Espinosa, R. Lin, Z. Liu, J. Padilla, and C. Delimitrou. "An Open-Source Benchmark Suite for Microservices and Their Hardware-Software Implications for Cloud & Edge Systems". In: *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems*. ASPLOS '19. Providence, RI, USA: Association for Computing Machinery, 2019, pp. 3–18.

[2] L. Qu, Q. Wang, T. Chen, K. Li, R. Zhang, X. Zhou, Q. Xu, Z. Yang, C. Yang, W. Qian, and A. Zhou. "Are current benchmarks adequate to evaluate distributed transactional databases?" In: *BenchCouncil Transactions on Benchmarks, Standards and Evaluations* 2.1 (2022), p. 100031. ISSN: 2772-4859. DOI: https://doi.org/10.1016/j.tbench.2022.100031.

TUDelft
Delft University of Technology