Stateful Functions on Streaming Dataflows

Kyriakos Psarakis

TU Delft, The Netherlands
k.psarakis@tudelft.nl

In order to migrate traditional monolithic applications to the cloud, developers typically adopt a microservice-like architecture. Although this migration provides substantial benefits such as scalability and development agility, it also leaves behind the transactional guarantees that database systems have provided to monolithic applications for decades. In the cloud era, developers build transactional and fault-tolerant distributed applications by explicitly programming transactional protocols at the application level.

The Stateful Functions-as-a-Service (SFaaS) paradigm aims to serve these use cases. However, existing approaches [1, 2, 4] provide weak transactional guarantees or perform expensive external state accesses requiring inefficient transactional protocols, increasing execution latency.

In our work, we argue that the principles behind the streaming dataflow execution model and deterministic transactional protocols provide a robust and suitable substrate for executing transactional cloud applications. To this end, we introduce Styx[3], a transactional application runtime based on streaming dataflows that executes serializable transactions consisting of stateful functions that form arbitrary call-graphs with exactly-once guarantees. Styx extends a deterministic transactional protocol by contributing: (i) a function acknowledgment scheme to determine transaction boundaries required in SFaaS workloads, (ii) a function-execution caching mechanism, and (iii) an early-commit reply mechanism that substantially reduces transaction execution latency. Experiments with the YCSB, TPC-C, and Deathstar benchmarks show that Styx outperforms state-of-the-art approaches by achieving at least one order of magnitude higher throughput while exhibiting near-linear scalability and low latency.

References

[1] Z. Jia and E. Witchel. Boki: Stateful serverless computing with shared logs. In *Proceedings of the ACM SIGOPS 28th Symposium on Operating Systems Principles*, pages 691–707, 2021.

[2] Q. Li, P. Kraft, K. Kaffes, A. Skiadopoulos, D. Kumar, J. Li, M. J. Cafarella, G. Graefe, J. Kepner, C. Kozyrakis, M. Stonebraker, L. Suresh, and M. Zaharia. A progress report on DBOS: A database-oriented operating system. In *CIDR*, 2022.

[3] K. Psarakis, G. Christodoulou, M. Frankoulis, and A. Katsifodimos. Transactional Cloud Applications Go with the (Data)Flow. In *CIDR*, 2025 ***(to appear)***.

[4] V. Sreekanti, C. Wu, X. C. Lin, J. Schleier-Smith, J. Gonzalez, J. M. Hellerstein, and A. Tumanov. Cloudburst: Stateful functions-as-a-service. *Proc. VLDB Endow.*, 2020.