NEWS LETTER

10

JANUARY | 2025

CLOUD STARS

Best Student Paper Award:

Bringing Function-as-a-Service to Voluntary Computing at the Edge
Catarina Gonçalves, José Simão and Luis Veiga

Enhancing decentralization, dependability and performance in cloud applications

by Luis Veiga (INESC-ID)

Cloud computing is the predominant paradigm today to provide computing resources, services, and applications. In particular, Function-as-a-Service (FaaS) has become a promising cloud computing service delivery model, well-suited for processing the massive data streams generated by the Internet of Things. Despite its potential, existing FaaS solutions face challenges in efficiently utilizing resources across distributed edge devices. Catarina Gonçalves, a student of the University of Lisbon carried out research at INESC-ID and published a paper about FaaS@Edge[1], a novel approach that leverages volunteered edge node resources, discovered via the IPFS network, to deploy functions using the Apache OpenWhisk framework. FaaS@Edge enhances system scalability and efficiency, with its evaluation results indicating that while FaaS@Edge introduces a slight latency overhead for function submission, its invocation times are comparable to local OpenWhisk deployments. Additionally, the system achieves high reliability, with an overall request success rate of approximately 98% for both submission and invocation. Her work has earned the Best Student Paper Award at the 20th International Conference on Economics of Grids, Clouds, Systems & Services (GECON 2024), in Rome, last September.

As more and more applications and systems are migrated to the cloud, and of increased mission importance, dependability becomes an important issue besides crash-fault-restart fault-tolerance. Significant research has been conducted on Byzantine Fault Tolerance (BFT) systems, leading to the development of novel techniques for addressing challenges such as operation ordering and state transfer in networks affected by Byzantine faults. More recently, the rise of distributed ledger technologies has renewed interest in BFT systems, driven by their superior throughput compared to other Byzantine consensus mechanisms. In collaboration with University of Oporto, INESC-ID is aiming to make this transition easier and more flexible by means of a Rust-based BFT middleware framework. Atlas [2] is an open-source, modular BFT framework designed to facilitate the development of high-performance BFT protocols. By decoupling traditionally intertwined components - such as separating the consensus primitive from execution and delegating log management to replicated services - Atlas enhances reusability and flexibility. It supports diverse BFT models, including deterministic and probabilistic/randomized approaches, through its modular design. The framework validation is based on the comparative ease of development of a novel BFT algorithm (FeBFT) with a number of optimizations, as well as on the evaluation and benchmarking of state-of-the-art BFT-SMR implementations. Ultimately, Atlas aims to: improve reusability and productivity in BFT development; enhance system safety, performance, scalability, and recovery efficiency through targeted optimizations; and foster reproducibility across BFT (sub-)protocols.

Finally, cloud computing has been a great enabler for Big-Data processing and the today's pervading AI workloads. However, these require increasing amounts of computing power and energy, with the risk of making them unsustainable w.r.t. cost and environmental impact. An ongoing approach to improve this is by attempting to perform computations as close as possible to the data (near-data processing), also referred as out-of-core computing. A promising avenue for this is by deploying in-memory, and even in-situ/in-storage compute devices. But as of now, this kind of device is less available and more costly than typical CPUs and traditional memory (volatile and non-volatile). Miguel Coimbra, a post-doctoral researcher at INESC-ID, has is just completing a secondment at IBM Research Europe - Zurich, where he explores how modern graph processing frameworks, such as Chunk Graph, and typical graph workloads could benefit from these kinds of devices in the near future, while assessing how their usage can be made more efficient, regarding cost-performance. His preliminary investigation into identifying I/O bottleneck patterns within out-of-core graph processing frameworks, with the motivation of potential application of custom storage hardware to alleviate these bottlenecks, has revealed that a small fraction of graph partitions residing in few blocks (5%) accounted for a significant majority of computation and I/O accesses (75%). Thus, it can be hypothesized that future cloud computing deployments could leverage specialized hardware for only a fraction of the workload data (such critical blocks) to perform out-of-core graph computing operations that could significantly improve execution performance with much reduced added cost.

[1] FaaS@Edge: Bringing Function-as-a-Service to Voluntary Computing at the Edge. Catarina Galvão Gonçalves, José Simão, Luís Veiga. 20th International Conference on the Economics of Grids, Clouds, Systems, and Services (GECON 2024), LNCS, Springer (Best Student Paper Award).

[2] Atlas, a modular and efficient open-source BFT framework. Nuno Neto, Rolando Martins, Luís Veiga. Journal of Systems and Software, 112317, DOI https://doi.org/10.1016/j.jss.2024.112317, available online Dec. 2024.

[3] CloudStars: Graph workload overheads and specialized hardware (WIP). Miguel Coimbra et. al. Jan 2025 (under development).

cloudstars.eu | twitter.com/Cloudstars_2023 | github.com/cloudstars-eu

CLOUDSTARS project has received funding from the European Union's Horizon research and innovation programme under grant agreement No 101086248