

The workflow motif: A powerful abstraction for debugging distributed applications

Mania Abdi, Northeastern University and MOC

Golsana Ghaemi, Boston University and MOC

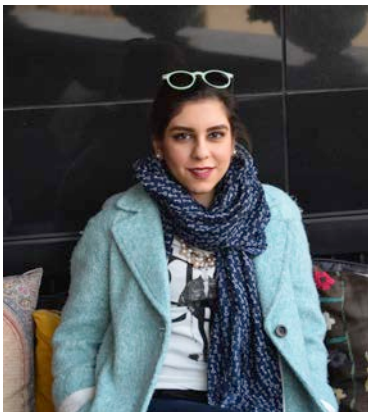
Abstract

Abstractions, such as APIs and libraries, enable complex distributed applications to be built. This is because they allow developers to build more complicated applications out of smaller building blocks (i.e., smaller applications) w/o necessarily understanding the latter's implementation details. But, when it comes to problem diagnosis, there are very few abstractions to help engineers. In many cases, engineers must diagnose problems using no abstractions whatsoever (i.e., using flat logs of raw application events). In this talk, we discuss the workflow motif, which we argue is a powerful abstraction that is useful for a variety of diagnosis tasks. Workflow motifs describe frequent processing patterns observed across distributed application requests. They represent the building blocks of how distributed applications process different requests. Understanding their properties provides significant insight into a distributed application's performance and correctness.

Speaker Bios



Mania Abdi is a PhD student at the Northeastern University Solid State Storage research group. Prior to that, she was a software engineer. She received her MSc. in Computer Engineering at the Sharif University of Technology and B.Eng. in Computer Engineering at the Amirkabir University of Technology. Mania is a computer systems researcher with a storage focus and has worked on a broad set of topics, including distributed storage, caching, data center debugging, and end-to-end tracing.



Golsana Ghaemi is a PhD student at Boston University cloud computing group who is working with Professor Orran Krieger and Dr. Raja Sambasivan. Prior to that, she was a computer architect researcher at Sharif University of Technology where she got her MSc. in Computer Engineering as well. Golsana is a computer systems researcher with a focus on end-to-end tracing, distributed systems, and cloud computing.