Project title: Leveraging ML to automatically detect performance issues in cloud applications

Brief Description of Design Project Goals:

Overview: Cloud applications are shifting from large monolithic applications that encompass all functionality, to complex graphs of hundreds or thousands of microservices, each corresponding to a small fraction of the end-to-end service.

A major challenge with microservices is detecting where performance issues initiate and which microservice is the culprit of a quality-of-service violation. Distributed tracing systems, like Zipkin or Dapper, can help in this direction by collecting end-to-end request traces.

The goal of this project is to apply ML and data mining techniques on distributed traces to automatically detect performance issues, and the microservices that initiated them.

Specific MEng Contribution: This project will involve the following tasks for an M.Eng student: (1) set up Zipkin on a few representative distributed applications that use microservices. (2) collect end-to-end traces using Zipkin. (3) develop ML/data mining techniques that automatically determine performance discrepancies, and pinpoint which microservices cause quality-of-service violations. The system will be evaluated on a local cluster of multicore servers. For any questions, contact Christina at: delimitrou@cornell.edu

ECE Field Advisor Name: Christina Delimitrou

- Email – delimitrou@cornell.edu
- Phone – 255-9316
- Office – 332 Rhodes Hall

Number of MEng Students Needed: 2

Required Skills: experience in C++/Python/Javascript, Linux, basic computer systems principles.

Estimated Project Time Frame:

2017-18 Academic Year, Two (2) Semesters