Tutorial Proposal for IC2E 2022

Tutorial Title: FaaSET: A Jupyter notebook to streamline serverless development, performance analysis, and experimentation

Tutorial Abstract (175-word limit)
Serverless Function-as-a-Service (FaaS) platforms automate many features of distributed software that are notoriously difficult to implement including high availability, fault tolerance, and automatic scaling while only billing for the actual runtime of functions in sub-second intervals. This tutorial will introduce the Function-as-a-Service Experiment Toolkit (FaaSET), a Jupyter notebook that helps to streamline serverless software development, performance analysis, and experimentation. We will provide an overview of existing Function-as-a-Service serverless platforms and discuss current challenges and limitations. We will then introduce FaaSET, which offers a development environment enabling developers to write functions, deploy them across multiple platforms, invoke and test them, automate experiments, and perform data analysis all in a single environment. FaaSET leverages the Serverless Application Analytics Framework (SAAF), an analytics framework that supports profiling workload performance, resource utilization, and infrastructure to enable performance and cost characterizations of serverless software deployments. The FaaSET notebook currently supports multiple FaaS platforms including AWS Lambda, Google Cloud Functions, IBM Cloud Functions, Azure Functions, and OpenFaaS.

Presenter Contact information
Robert Cordingly
rcording@uw.edu
School of Engineering and Technology
University of Washington Tacoma

Wes J. Lloyd
wlloyd@uw.edu
School of Engineering and Technology
University of Washington Tacoma

Robert Cordingly - Brief Biography
Robert Cordingly is a PhD student and research assistant at the University of Washington Tacoma in the Computer Science and Systems program. His research focuses on serverless cloud computing, cloud performance analysis and cost modeling, and serverless application design. He is the primary developer for FaaSET, SAAF, FaaS Runner, and other supporting tools. With his tools and research, he aims to aid developers and researchers in making educated choices to improve application performance and cost on cloud computing platforms.

Wes Lloyd - Brief Biography
Wes Lloyd is an assistant professor in the Computer Science and Systems program at the School of Engineering and Technology at the University of Washington Tacoma. Wes teaches

classes in cloud computing and systems. His research interests span cloud computing, distributed systems, and software engineering. His research aims to help scientists and practitioners better leverage cloud computing technologies to improve performance, availability, and observability of software deployments to maximize efficiency and lower hosting costs.

Proposed Duration

Half-day (3 hours): The tutorial is proposed as a half-day tutorial to encourage and maximize participation. Longer tutorials involve a higher time commitment for participants and are more likely to suffer from participant attrition.

Tutorial History

This tutorial was previously presented as an online tutorial at the 2021 ACM/SPEC International Conference on Performance Engineering (ICPE 2021). This instance of the tutorial will focus on the FaaSET Jupyter notebook, a new tool leveraging SAAF created since ICPE 2021. FaaSET provides a development environment enabling developers to write functions, deploy them across multiple platforms, invoke and test them, automate experiments, and perform data analysis all in a single environment

Intended Audience Skill Level

This tutorial will be offered at the beginner/intermediate level. Previous experience programming in Python is preferred.

Tutorial Detailed Description (3-page limit)

Upon successful completion of this tutorial, the participants will be able to:

- Understand advantages and challenges of developing software leveraging cloud-based serverless Function-as-a-Service (FaaS) platforms
- Understand the capabilities of the Serverless Application Analytics Framework (SAAF) including the performance metrics it provides to help characterize performance, resource utilization, and infrastructure of software deployments to FaaS platforms
- Write a serverless FaaS function on a public cloud platform and integrate source code with SAAF to enable performance evaluation
- Understand and use the FaaSET Jupyter notebook which integrates theFaaS Runner tool to enable repeatable performance experiments and aggregate metrics to characterize performance, resource utilization, and infrastructure

Draft Schedule of the Tutorial (Half day)

Our tutorial consists of both lectures and hands-on exercises.

Logistics: Prior to the tutorial we will share detailed instructions to guide participants to obtain access to a public cloud Function-as-a-Service platform for the purpose of participating in the tutorial. All of the tutorial slides and handouts will be made available to the participants prior to the tutorial. At the end of the tutorial, we will provide references, URLs, GitHub repositories, and additional resources so that participants can continue to explore and leverage SAAF and its related tools.

This tutorial will be composed of a lecture introducing serverless computing, FaaS, SAAF, the FaaSET notebook followed by a demonstration and a hands-on interactive tutorial. During the hands-on activities, we expect to have at least two presenters available to answer questions from the participants. We will provide a survey to collect user feedback at the end of the tutorial.

First Half - Lecture and Demo (90 minutes):

- *Motivation and background*
    - Introduction to serverless computing and Function-as-a-Service (FaaS) platforms
    - Advantages and challenges of FaaS platforms
    - Comparison of existing public cloud and open source FaaS platforms
- Introduction of the Serverless Application Analytics Framework (SAAF)
    - Performance, resource utilization, and infrastructure metrics available from SAAF
    - Review of supported platforms and languages
    - Demonstration on how to integrate SAAF with new and existing serverless functions
- **DEMO:** Introduction of the FaaSET notebook incorporating tools: SAAF and FaaS Runner
    - Creating serverless functions in the FaaSET notebook
    - Deploying functions to multiple FaaS platforms with FaaSET
    - Demonstration of interactive reproducible experiments with FaaSET notebook
    - Scalability testing with FaaSET: how to automate parallel and scalable workloads to evaluate performance
    - Memory testing with FaaSET: how to test function performance over various memory configurations
    - Performance analysis utilizing SAAF metrics and generate visualizations

Second Half – Hands-On Activities (90 minutes):

- **ACTIVITY**: Creating FaaS functions using the FaaSET notebook
- **ACTIVITY**: Automating multi-threaded concurrent workloads for scalability testing
- **ACTIVITY**: Automating memory testing to find the ideal memory reservation size to balance performance and cost
- Introduce Performance Analysis and Graphing with FaaSET
    - Overview of available features and data aggregation capabilities
    - Example of loading FaaS Runner experiment results and data processing
    - Utilizing Jupyter notebook libraries (e.g. Plotly, Matplotlib) to generate and visualize graphs
- Discussion: We will wrap up the tutorial by having an open discussion to solicit feedback from the participants on the tools and educational materials. Participants will be encouraged to suggest desired features.

Tutorial survey: At the conclusion an online Google Forms survey will be used to elicit digital feedback from tutorial participants to augment the discussion in the event of limited time.

Tutorial Relevance for IC2E

This tutorial presents the Serverless Application Analytics Framework (SAAF), a framework designed to enable **performance measurement** and **performance modeling** of software deployments made to serverless Function-as-a-Service (FaaS) platforms. The SAAF framework characterizes the performance, resource utilization, and virtual infrastructure utilization of software deployments to FaaS platforms improving observability to enable performance engineering. As cloud computing service delivery models evolve to embrace the "serverless" paradigm, observability of factors that underpin performance of software deployments has become increasingly inaccessible to developers. Tools such as SAAF are paramount to improve observability of serverless software deployments. SAAF captures numerous metrics to provide insights to support root cause analysis of performance outcomes enabling developers to better understand and optimize for software design tradeoffs that underlie serverless platforms. This tutorial also showcases the interactive FaaSET Jupyter notebook, an integrated development environment enabling developers to write functions, deploy them across multiple platforms, invoke and test them, automate experiments, and perform data analysis all in a single environment

Related Terms from IC2E Call for Contributions for this tutorial:

Cloud management and engineering

- Resource management and optimization
- Service lifecycle management, automation
- Performance, dependability, SLAs, Quality of Service
- Deployment and migration
- Monitoring, benchmarking, and testing
- Software Engineering methods for cloud services and applications
- Programming models and tools
- Microservices
- Operational analytics and DevOps solutions

Cloud applications

- Web and enterprise computing
- Cloud systems

Cloud storage and databases, data management

- Middleware platforms
- Serverless computing, Function-as-a-Service
- Platforms and services for data and analytics

Audio-visual and Technical Requirements

We intend to present the tutorial as a hybrid in-person / online interactive session. We can host the session over Zoom or use the video conferencing platform provided by IC2E. Tutorial participants will access one or more public clouds for activities during the tutorial. Registered participants will be provided with instructions for how to establish FREE access to a public cloud

platform for use in the tutorial. Participants that already have public cloud computing accounts will be able to use their existing accounts with no setup required prior to the tutorial.

During the tutorial we will share detailed instructions including screen captures to guide participants through hands-on exercises. All tutorial slides and handouts will be made available to the participants prior to the tutorial. At the end of the tutorial, we will provide references, URLs, GitHub repositories, and additional resources to all participants to support further exploration and use of SAAF and our function-as-a-service experiment toolkit, FaaSET.

References

Cordingly, R., Lloyd, W., FaaSET: A Jupyter notebook to streamline every facet of serverless development, 2022 13th ACM/SPEC International Conference on Performance Engineering: 5th Workshop on Hot Topics in Cloud Computing Performance (HotCloudPerf-2022), Apr 9, 2022.

Cordingly, R., Heydari, N., Yu, H., Hoang, V., Sadeghi, Z., Lloyd, W., Enhancing Observability of Serverless Computing with the Serverless Application Analytics Framework, Tutorial Paper. 2021 12th ACM/SPEC International Conference on Performance Engineering (ICPE '21), Apr 19-23, 2021 (tutorial paper).

Cordingly, R., Hanfei Y., Hoang, V. Sadeghi, Z., Foster, D., Perez, D., Hatchett, R., Lloyd, W, The Serverless Application Analytics Framework: Enabling Design Trade-off Evaluation for Serverless Software. In Proceedings of the 2020 Sixth International Workshop on Serverless Computing, pp. 67-72, December 2020.

Cordingly, R., Shu, W., Lloyd, W., Predicting Performance and Cost of Serverless Computing Functions with SAAF, In Proceedings of the 2020 Sixth International IEEE Conference on Cloud and Big Data Computing, (CBDCom 2020), pp. 640-649, August 2020.

Cordingly, R., Hanfei Y., Hoang, V., Perez, D., Foster, D., Sadeghi, Z., Hatchett, R., Lloyd, W., Implications of Programming Language Selection for Serverless Data Processing Pipelines, In Proceedings of the 2020 Sixth International IEEE Conference on Cloud and Big Data Computing, (CBDCom 2020), pp. 704-711, August 2020.