**Proposal document**

Visualization of system performance on FECBench framework

Wei Fan

**Project link**

<https://github.com/FWWorks/VAML_ApplicationPerformance>

**Basic information**

Performance interference caused by contention for non-partitionable resources in multi-tenant platforms always lead to unpredictable behavior and degradation in application performance. To address this problem, many methods or tools have been proposed and developed such as iBench, DIAL, and FECBench. In these methods or tools, performance analysis plays an important role on monitoring resource stress, evaluating performance, and predicting performance abnormality etc. To gain insight into performance analysis, visualization of performance analysis, which is a type of software visualization that includes aspects such as hardware performance, computation/system visualization, should provide abundant of both intuitive and interactive visualizations presenting what the interference is, whether the behavior is predictable and how the application performance is.

My term project is aimed to develop a visualization tool for one of the interference-aware methods for application performance modeling – FECBench.

**Background**

FECBench (Fog/Edge/Cloud Benchmarking) is an open source framework comprising a set of 106 applications covering a wide range of application classes that guides providers in building performance interference prediction models for their services without incurring undue costs and efforts via the following contributions [1].

FECBench comprises an ofﬂine stage with a set of steps to create a knowledge base followed by an online stage (See Figure (a)). Developers can use the same ofﬂine stage process to further reﬁne this knowledge base. Accordingly, the ﬁrst step of the ofﬂine stage deﬁnes a Benchmark Warehouse (BMW), which is a collection of resource utilization metrics obtained by executing a large number and variety of applications. The second step clusters these applications according to their similarity in how they stress individual resources. Then the knowledge base captures the stress on resources stemming from executing a combination of colocation patterns of applications belonging to the different clusters. After that, FECBench deﬁnes a resource stressor prediction model and uses design of experiment to give rise to a knowledge base. FECBench builds Application Mix Repository that cover the Resources Design Space Surface [1].

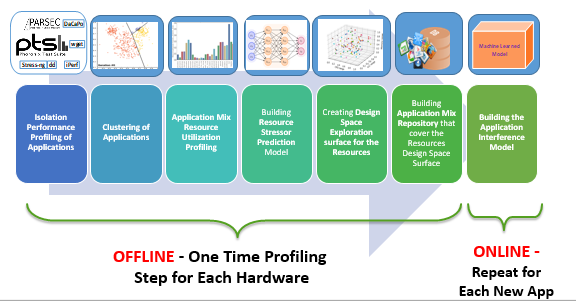


Figure (a) FECBench workflow

**Project Description**

In my project, I want to visualize dynamic execution behavior of the target application’s performance analysis and profiling steps in the process of generating design space described in the background section.

To be more specific, my visualization tool should provide dynamic monitoring visualizations since the FECBench is aimed to monitor the performance of target applications. For example, I need to plot the overall application performance with time series. Overall application performance is a scalar value which refers to quality of service (QoS) for Internet of Things (IoT) applications, and response time for web applications etc. I want to plot the application performance along with the simple analysis of resource stress system suffering and the resource stress this target application brings about. These analysis are important to give user an overall insight on ‘How is my application running?’, ‘Which system metric is the dominant in my application? ’, ‘How can my application run in such resource stress?’ etc.

Since FECBench provides prediction of application performance, anomaly of performance should also be an important aspect to visualize. For example, for Internet of Things applications, it is necessary to ensure that a topic’s latency is within a desirable quality of service value. So if an iot application cannot reach the desirable quality of service value, visualization should be able to present that as an anomaly situation. And maybe some plots about history performance of the application can be presented too. Relevant applications which has similar latency criterion can also be shown for comparison. And user can tune the desirable quality of service value in an interactive level.

As for the profiling offline stages in FECBench framework, I want to plot the benchmark warehouse, which is the first step of the six stages. Then I want to do visualization for clusters generated from the second offline stage and add some interactive methods such as allowing user to choose metrics for clustering to show how clusters stress individual resources. Comparison between resources profiling of different clusters should also be visualized to show the difference or similarities. Because FECBench uses design of experiment to give rise to a knowledge base, I want to make visualizations for knowledge base or resource stressor prediction model to make visual analysis for model training and try to create visualization for the design space. For that design space is multi-dimensional, reduction dimension methods may be used.

**Data**

For the dynamic monitoring visualizations, I need different system metrics data which can compose a design space and a good scalar to describe system stress, such as Cache Hit/Miss, Disk IO, Network IO, L3 cache Bandwidth, Memory Bandwidth, CPU Utilization, IPC etc.

Also, time series of benchmark performance data on the system performance metrics are needed to complete the prediction task.

Since I want to do visualization of the six offline stages of FECBench framework, I need the 106 applications’ data of stress on resource to do cluster comparison and present the resource stressor prediction model.

For scaling, I need some criterion data to give overall insight to evaluate whether the performance is abnormal.

**Visualization library**

To get an initial sight into raw data, I would like to use some high-level authoring tools like plotly or vega-lite.

And to implement the dynamic monitoring visualizations and profiling, I will use D3.js and React. If I find a 3D plot will be better, I will use Three.

**Baseline**

In order to do visualization on resource stressor, which uses multiple dimensions (>3) to represent, I need to use PCA to reduce dimension. Therefore, I want to apply iPCA to visually analyze system metric data and perform interactive actions. [Paper: iPCA: An Interactive System for PCA-based Visual Analytics]

Regarding to visualizing dynamic execution behavior of the target application’s performance analysis, I want to follow some ideas like the integrated monitor and response time heat map from the application performance monitor tool of solarwinds. I also want to use some ideas from paper ‘Memory Heat Map: Anomaly detection in real-time embedded systems using memory behavior’, which uses heat map to do anomaly detection, to implement a resource stress heat map or a performance heat map to reflect anomaly.

For cluster analysis, I want to follow ideas from paper ‘Clustrophile 2: Guided Visual Clustering Analysis’ to give the resource utilization metrics of a variety of applications clusters a data table that supports a rich set of filtering interactions and statistics and multiple resizable Clustering Views, which can be used to work simultaneously on different clustering instances and provides several ways to guide users during their analysis.

**Schedule**

Review and clean the data (1 week)

Learn how to use the library or tools (1-1.5 weeks)

Set up project frame (2 weeks)

Implement static visualization (1 week)

Implement dynamic monitoring visualization (1.5 weeks)

Implement cluster analysis visualization (1.5 weeks)

Implement anomaly prediction visualization (1 week)

Implement new ideas or some supplementary visualization (1 – 2 weeks)

**References**

[1] FECBench: A Holistic Interference-aware Approach for Application Performance Modeling