Facilitating continuous benchmarking for the critical components of SkyhookDM

Rahul Agrawal
(Mentored by Jayjeet Chakraborty)
UC Santa Cruz
SkyhookDM

- Programmable object storage built on Ceph distributed object storage.
- Supports database operations such as SELECT, PROJECT, AGGREGATE to be offloaded (i.e., pushed down) to the object storage layer
- Partitions tabular data; stores partitions in objects
- Methods are implemented to be applied by objects to their data locally
Problem

- SkyhookDM is a performance-critical computational storage system developed by embedding Apache Arrow inside the Ceph storage layer.
- Small changes in the source code’s performance-critical parts will often result in significant performance changes.
- It’s essential to keep track of these performance changes so that the project can become more performant over time and silent performance degradation can be avoided.
Solution

- A benchmarking framework can be used to create benchmarks (very similar to unit tests) for all the performance-critical parts of the source code.
- These benchmarks can be added as a separate job in the CI/CD pipeline, which get triggered when any particular events like commit/push happen.
- Some visuals can also be integrated to monitor the performance results of the CI tests.
Phase 1 - Integrated Conbench

- Conbench is a language-independent, continuous benchmarking framework, built specifically with the needs of a cross-language, platform-independent, high-performance project like Apache Arrow in mind.
- It is successfully integrated with all the necessary changes in the codebase.
- Some additional benchmark tests are written to test every aspect of the SkyhookDM
Phase 2 - Automated Benchmarking using Github Actions

- All the benchmarks are now executed automatically as soon as a new pull request is created
Phase 3 - Visualizing the Outputs

- Python scripts added to visualize the outputs in a better way
Phase 4 - Comment Bot Integrated

- Integrated a comment bot that adds a comment (containing the Plots) in the same PR thread
Thank You!

Questions?

rahulagrawal799110@gmail.com