Optimising the pipelines using Github Actions for Caliper and Caliper-Benchmarks

Project Title	Optimising the pipelines using Github Actions for Caliper and Caliper-Benchmarks
Status	COMPLETED
Difficulty	MEDIUM

Description

Performance benchmarking distributed systems (such as DLTs) is a challenging task, comprising of multiple aspects, such as scalable workload generation, representative workload definitions, and comprehensive data analysis.

Hyperledger Caliper is a general-purpose benchmarking tool with the goal of mitigating the aforementioned aspects of performance benchmarking:

It provides a flexible architecture to allow scalable workload generation.
It collects and reports results based on detailed client-observable execution traces.
It allows the implementation/plug-in of custom workload behaviors to meet the diverse criteria of a wide range of business scenarios.

Hyperledger Caliper-benchmarks provides a set of example benchmarks for use with Hyperledger Caliper.

In order to be able to assist maintainers/contributors to Caliper and Caliper-Benchmarks, a build system is utilised to run tests to verify a PR doesn't break anything plus it provides a continual deploy mechanism to publish Hyperledger Caliper artifacts to DockerHub and NPM so end users can have access to the Caliper tool.

These build systems use different build pipeline implementations. Hyperledger Caliper uses Azure Pipelines and Hyperledger Caliper-Benchmarks uses Github Actions. Azure pipelines is going to be removed in the future so we need to ensure both projects are using Github Actions. Further these build systems run all available tests which is not necessary and thus uses more energy than is needed and also means that these builds take longer meaning a PR cannot be merged until this build has completed (or may require multiple runs based on PR reviews). This is not ideal. Further although artifacts are published externally, no testing is done to ensure that what would be published actually works which means it's possible to deliver something to users which doesn't work.

The goal of this project is to optimise how Caliper and Caliper-Benchmarks verify PRs and publish artifacts using Github Actions and to improve how testing is done to ensure that the artifacts to be published actually work as expected.

The exact steps of the project are detailed in the Expected Outcome section.

Additional Information

- Caliper GitHub repository: https://github.com/hyperledger/caliper
- Caliper example benchmarks GitHub repository: https://github.com/hyperledger/caliper-benchmarks
- Caliper documentation page: https://hyperledger.github.io/caliper/
- Architecture documentation: https://hyperledger.github.io/caliper/vNext/architecture/
- Benchmark configuration documentation: https://hyperledger.github.io/caliper/vNext/bench-config/
- Workload module documentation: https://hyperledger.github.io/caliper/vNext/workload-module/
- Discord: https://discord.gg/hyperledger

Learning Objectives

- Mastering the open-source workflow
  - Issues, pull requests, development branches, repository management
- Being part of an open community, navigating and utilizing different forums
- Working in a team
- Covering a wide spectrum of software development skills
  - Conducting unsupervised research for state-of-the-art solutions
  - Writing detailed, consistent specifications
  - Designing/implementing functionally rich, testable, maintainable software components
  - Documentation and presentation skills
    - Developer documentation
    - To-the-point, coherent presentation

Expected Outcome

1. Migrate Caliper from Azure Pipelines to Github Actions

2. Investigate and propose how to optimise the builds to reduce PR/Publish build times and save resources and energy, some thoughts on possible suggestions could be:
    a. only runnning tests which will test affected code changes
    b. caching of docker images and npm modules
3. Implement the above proposal
4. Investigate and propose how to verify publishable artifacts before they are published, some thoughts here are
    a. testing the docker image
    b. testing the npm modules
5. Implement the above proposal

Relation to Hyperledger

Hyperledger Caliper and Hyperledger Caliper-Benchmarks

Education Level

At least ongoing B.Sc. studies in software engineering required.

Skills

- Basic understanding of version control and git
- Experience with Bash
- Experience with JavaScript programming and npm (in Node.JS context)
- Experience using Docker
- Intermediate verbal and writing skills in English
- High-level understanding of Hyperledger Caliper components and the benchmark workflow

Nice-to-have skills (the mentee can learn these during the internship):

- Experience with github actions
- Advanced git usage (upstream repositories, branching, rebasing, etc)
- Advanced NPM and Docker usage
- Writing documentation in markdown
- Capability for unsupervised learning and research

Future plans

As new features are added to Caliper and Caliper Benchmarks, so the build system needs to evolve to ensure that the builds remain current, provide what's needed for PR verification and artifact publishing as well as ensuring optimal use of resources to keep energy consumption down

Preferred Hours and Length of Internship

Full-time (40 hours a week for 12 weeks during the summer) or Part-time (20 hours a week for 24 weeks starting in summer and ending in fall)

Mentor(s) Names and Contact Info

David Kelsey, d_kelsey@uk.ibm.com, discord ID: NorfolkAndChance#6513