TP3: Comparison between NoSQL and Blockchain as distributed databases.

This assignment focuses on the subject of Software Architectures for Big Data. As discussed in class, Blockchain can be perceived as a distributed database. The main question of the assignment is how a Blockchain system compares against a regular NoSQL database with respect to performance, if we ignore the security advantages. In this assignment, you are tasked with writing an academic paper to present a comprehensive comparison between Hyperledger Fabric and a NoSQL database of your choice. More specifically, you have the following tasks:

1. Provide a high-level comparison of Hyperledger Fabric and the NoSQL database with respect to their documented properties and the general attributes of Big Data architectures.
2. Use the tool Hyperledger Caliper (https://github.com/hyperledger/caliper) to benchmark the performance of Hyperledger Fabric with respect to read/transaction latency and throughput and resource consumption metrics (CPU, memory, network IO).
3. Use the Yahoo! Cloud Service Benchmark (https://github.com/brianfrankcooper/YCSB) to benchmark the performance of your NoSQL database. Deploy an instance of your selected implementation and test its performance using the workloads provided by the benchmark. Similarly, to Hyperledger Fabric provide read/write latency and throughput and resource consumption metrics. Writes in NoSQL are considered equivalent to transactions in Blockchain.
4. Provide a comprehensive comparative evaluation for all three implementations and discuss the differences in the performance.
5. You can use the workload mixes specified in YCSB (only read, 50/50 read/write, 10/90 read/write) to specify similar workloads in Caliper as well.

To further help you with the assignment, here is an expected structure of the paper:

1. Abstract: Summarize your work in no more than 150 words.
2. Introduction
    a. What is the main motivation behind using distributed databases?
    b. What are their main features and characteristics?
    c. What are the objectives and contributions of this work?
3. Background
    a. Present a brief description of Hyperledger Fabric with its main characteristics and everything that is relevant for your paper.
    b. Present a brief description of your NoSQL DB with its main characteristics and everything that is relevant for your paper.
    c. Compare the two based on their documentations and basic properties
4. Experiments
    a. Describe your experimental setup
        i. Deployment for Hyperledger Fabric (infrastructure, configuration etc.)
        ii. Deployment for the NoSQL DB (infrastructure, configuration etc.)
        iii. Description of the employed workloads
        iv. Setup of Hyperledger Caliper and YCSB

    b. Results
        i. Present tables and charts for the performance of your two systems.
        ii. Describe the results.
5. Discussion
    a. Compare the results of your systems and justify their differences.
    b. Did your work have any limitations or threats to validity?
6. Conclusions
    a. Summarize your findings and present any possible future work.

Submission Guidelines

- Submit a paper in IEEE format (https://www.ieee.org/conferences/publishing/templates.html) of at least 8 pages (maximum 10 pages).
- Your paper should definitely have an abstract, keywords, an Introduction and conclusions. You can organize the rest of the content as you see fit.
- Provide a meaningful title (not just "Assignment 3") and give the author names and affiliations as specified by the template.
- This is an academic paper and you will need to consult numerous sources (other papers, documentation, possibly online posts). Make sure you properly cite your sources and give credit. Anything that does not have a citation will be considered your contribution and will be subjected to judgement. So, try to support as many of your arguments as possible with proper sources. DO NOT CITE WIKIPEDIA! Instead you can consult the references that a Wikipedia page already cites. They usually contain the information you are looking for. If you provide online documentation or blog posts (as a last resort), also provide the date you last accessed the source.
- Allocate enough time to set up and run the benchmark experiments with YCSB.
- Along with the paper, you will need to prepare and submit a presentation on your paper. The presentation will be given on the 3rd of December during regular class time by all team members. Each team will be allocated 12 minutes for the presentation and 2-3 minutes for questions.
- After the presentation you will have another week to submit the final version of your paper on the 9th of December. You may have to take into account feedback from the presentations to finish your paper.
- You will receive 70% for the paper and 30% for the presentation.
- **Presentation submission deadline: December 1st 23:59**
- **Paper submission deadline: December 9th 23:59**