Data Intensive Computing - Review Questions 1

Deadline: September 8, 2019

1. Explain how a file region can be left in an inconsistent state in GFS?

   _____

2. Briefly explain how HopsFS uses User Defined Partitioning (UDP) and Distributed Aware Transaction (DAT) to improve the system performance (e.g., CPU usage, network traffic and latency)?

   _____

3. Show with an example that in the CAP theorem, we can have only two properties out of Consistency, Availability, and Partition Tolerance at the same time.

   _____

4. How does BigTable provide strong consistency?

   _____

5. Write a simple query in Cypher for each of the following requests:

   - Match a Person called "John Doe"
   - Find FRIENDS_OF "John Doe"
   - Count "John Doe"'s direct acquaintances