ourdiu.com | Shared by MD. GOLAM MUSTAKIM NAEEM

Daffodil International University

Department of Computing and Information Systems

Course: Big Data Analysis and Applications — **Course Code:** CIS 315

Class Test 02 — **Time:** 30 Mins — **Total Marks:** 15 — **SET – A**

1. (a) Explain how *Velocity* and *Variety* present distinct architectural processing challenges compared to simple scaling of pure *Volume* in streaming analytics platforms. [3]
   (b) Distinguish between *Veracity* and *Valence* in Big Data ecosystems, providing a brief real-world example for each. [2]
2. (a) Why do traditional POSIX file systems fail when scaled across petabyte-level storage clusters? Detail the limitations of file locking and centralized metadata overhead. [3]
   (b) According to the CAP Theorem, explain why distributed network storage systems like HDFS must choose between Consistency or Availability during a network partition. [2]
3. (a) Describe the core responsibilities of the HDFS NameNode and explain how the "Small Files Problem" severely impacts NameNode RAM capacity. [3]
   (b) Outline the step-by-step data lifecycle during an HDFS Write Pipeline from the client stream to DataNode replication. [2]