**Student Name:**

**Paper 1 [ArrayStore]**

ArrayStore is a storage manager for complex, parallel array processing.

Q1: The paper presented 4 types of array chunking. List these 4 types and state why do we need to chunk a given array?

Q2: Assume we have an array in 2D space (X and Y) and each dimension is an integer between 0…1000. We partition the X-axis by ranges [0, 10], [11, 100], [101, 1000], and we partition the Y-axis by ranges [0, 500], [501, 1000]. What is the type of this partitioning or chunking?

Q3: ArrayStore proposed several methods to distribute the array chunks across the cluster nodes. Mention 4 types of these methods. Also, referring to Q2, can each array be distributed in a different way, or they have to follow the same distribution method? Explain why?

**Paper 2 [Stubby]**

Stubby system proposes several techniques to optimize MapReduce workflows. These techniques are based on transformations and generation of larger pool of query plans.

**Q1:** Is Stubby a cost-based system? Is it a rule-based system? Answer Yes/No and then explain your answer in one or two sentences. Also state what is the mechanism Stubby uses to carry the needed metadata information.

Q2: Stubby contains five types of transformations. For each of the following figures, state the type (name) of the transformation being performed, and write 2 or 3 sentences on what it does?

 

 

 

 