* Only relevant answers are graded.

01

a) When you consider: i) recursive and ii) immediate data decompositions? Explain.

b) How distributed memory computations benefit for task interaction graph? Give an example.

- c) What benefits are there in keeping two file chunks on the same rack in HDFS?
- d) How Namenode ensure accurate record of data-blocks? Explain with the assumption that there are at least 2-3 node failures per minutes in a 1500 node cluster.

Q2

```
What this program is doing? Explain the purpose of both if statement and each MPI ????? calls.
         int array_size = 10;
1
2
         int elem = array size / size;
3 🗸
         if (array_size % size == 0) {
4
             int* local_array, * global_array; // assume dynamic allocation and free
5 🗸
             if (rank == 0) {
                  for (int i = 1; i < size; i++) {
6
7
                      MPI_Send(&global_array[i * elem], elem, MPI_INT, i, 0, MPI_COMM_WORLD);
9
                 for (int i = 0; i < elem; i++) local_array[i] = global_array[i];</pre>
10 ~
             } else {
                  // Receive the portion of the array allocated to this node.
11
12
                 MPI_Recv(local_array, elements_per_node, MPI_INT, 0, 0, MPI_COMM_WORLD, MPI_STATUS_IGNORE);
13
14
15
         MPI_Finalize();
```

