Chennai Mathematical Institute

DISTRIBUTED COMPUTING AND BIG DATA   DURATION: 90 MINS. MAX
MARKS: 20.

ROLL NO.: _____   DATE: 02/06/2020

NAME: _____

Instructions

- Submit a single pdf file carrying your answers on moodle under "Mid Term" assignment. For any reason, if you cannot upload to moodle, email your work to vvtesh.cmi@gmail.com.
- A penalty of 1 mark applies for every minute of late submission (beyond 13:40 Hrs).
- This is an individual assessment. Do not discuss with anyone.
- This paper refers to a variable z. If the last digit in your roll number is i, then $z = (i\%4) + 1$ where $\%$ is the modulo operator. For example, if the last digit is 1, then $z = 2$.

> **Section 1: Questions 1 to 5 carry 3 marks each.**

Question 1. Ramesh has a file of size $4z$ terabytes. Ramesh wishes to send this file to Ria. He can send the file over a 400 Mbps direct dedicated network connection.

(1) How much time will it take for Ria to receive the file through the network?
(2) What is the maximum file size for which Ramesh will prefer to use the dedicated network channel over a overnight (i.e., within 24 hours) courier?

Explain your answer in detail. Include relevant calculations.

Question 2. Ram bought a new hard disk and configured it so that the number of bytes per inode is r. Prem too bought a new hard disk of same size and configured it to have the number of bytes per inode as p. Ram and Prem stored f number of files each in their respective disks. Given that $r > p$ and Ram had relatively smaller sized files, is it possible that Ram ran out of disk space while Prem did not? Explain with an example.

Question 3. A drive spins at 4800RPM and has average seek time of 12ms. The disk has $20 + 4z$ sectors per track. What is the average access time?

Question 4. Assume a disk size of $4z$ TB with block size of 8 KB.

(1) How much space will you need (in MB) to store the free space bitmap?
(2) If this free space bitmap needs to store additional information on whether the block is corrupt or not, how can you do it? How much space will you need to store this extended free space bitmap?

Question 5.   If we have 100 processors and 5% of the total jobs cannot be parallelized, what is the scaled speedup achievable as per Gustafson's law?

Section 2: Question 6 carries 5 marks.

Question 6.   You are provided with the following facts about a model of execution of a distributed system that uses global vector time stamps.
(1) $(1,1,1,1) \to (2,1,1,1)$ is a happens-before relation.
(2) Exactly two events occurred in each process.
(3) Between the two events of every process, at least one event occurred in another process.
(4) The first event occurred in the process p_3.

Agreeing to the above facts:
(1) Draw the space-time execution diagram annotated with global vector time stamps. (1.5 Marks)
(2) Draw the corresponding hasse diagram. (2 Marks)
(3) Draw the same space-time execution diagram annotated with matrix time. (1.5 Marks)