MANIPAL INSTITUTE OF TECHNOLOGY

MANIPAL

(A constituent unit of MAHE, Manipal)

I SEMESTER M.TECH. (COMPUTER SCIENCE AND ENGINEERING)

END SEMESTER EXAMINATIONS, MARCH 2021

HIGH PERFORMANCE COMPUTING SYSTEMS [CSE 5154]

REVISED CREDIT SYSTEM

Date: 01/03/2021

Time: 3 Hours MAX. MARKS: 50

Instructions:

Answer all the questions & missing data may be suitably assumed.

1A) Discuss any one of the operating system approaches to achieve parallel processing in a uniprocessor system. Draw the relevant timing diagram. **5M**

1B) Consider the Reservation Table of a pipeline shown in Fig. Q 1B.

	1	2	3	4	5	6	7	8	9
S1	X								X
S2		X	X					X	
S3				X					
S4					X	X			
S5							X	X	

Fig. Q 1B

a) Determine latencies in the forbidden list, *F* and the collision vector, *C*.
b) Draw the state transition diagram.
c) List all Latency cycles, Simple cycles and Greedy cycles.
d) Determine Minimal Average Latency.
e) Find out the efficiency η of this pipeline due to which pipeline can give best performance. **0.5M+2M+ 0.5M+0.5M+1.5M = 5M**

2A) With neat diagram, discuss the features and problems in crossbar switch network system for multiprocessors. **4M**

2B) Draw a diagram of 5D hypercube network. In the network, label your nodes in binary. Identify the node degree, network diameter and channel bisection width. **4M**

2C) For what purpose the Bus arbitration algorithms are used? Explain with diagram, how Polling technique works as one of the Bus arbitration algorithm. **2M**

3A) Write a MPI program to read a value *M* and *N* X *M* number of elements in the root. Using *N* processes do the following task. Find the square of first *M* numbers, find the cube of next *M* numbers and so on. Print the results in the root. Do not use any built-in math functions and also do not use point to point communication APIs. **5M**

3B) Write a parallel algorithm for 5D hypercube network which you have drawn in Q 2B, to sum *n* values where *n* is the number of values to be added and *p* is the number of processors in the model. It is assumed that the each processor initially holds one of the values among *n.* What is the time complexity that your algorithm has offered? **5M**

4A) Write an OpenCL kernel that takes a string S as input and one integer value N. Produce a string N times as follows in parallel:

Input: S = Kalam N = 4

Output String: KalamKalamKalamKalam

Note that each work-item copies same character from the Input N times to the required position in output string. Explain how the output string is generated parallelly by your kernel. **5M**

4B) How do you find out the execution time taken by a OpenCL kernel function starting from *event* object creation. Specify all the statements required for this. **3M**

4C) Explain workgroup and work item concepts by taking the example of a vector of 16 workgroups and 32 work items in each workgroup. **2M**

5A) Write a CUDA program to read a matrix of size 3 X 3. The kernel you write should meet the following specifications:

With respect to each row in the matrix, find first element to the power of 2, second element to the power of 3, third element to the power of 4. An example data is shown in Fig. Q 5A. You need to use number of CUDA threads as number of rows of matrix to make this task feasible. **5M**

Matrix A =	1	2	3	Matrix B =	1	8	81
	4	3	2		16	27	16
	3	2	2		9	8	16

Fig. Q 5A

5B) With block-thread organization in CUDA, you have to create 6 blocks in a grid and every block should contain 8 threads (i.e thread with index 0 through 7). Draw a neat block diagram to show how these blocks and threads are organized in 1D Grid of 3D Block. Label each and every blocks and threads of any one of blocks. Also write down the equations associated in computing the threadID in 1D Grid of 3D Block. **3M**

5C) Write down your conclusion with step number 10 of OpenCL implementation, “Configure the work-items structure” that works with the need of Q 5B. **2M**