U.S.N.										

B.M.S. College of Engineering, Bengaluru-560019

Autonomous Institute Affiliated to VTU

October 2024 Supplementary Examinations

Programme: B.E. **Semester: VI**

Branch: Artificial Intelligence and Machine Learning **Duration: 3 hrs.**

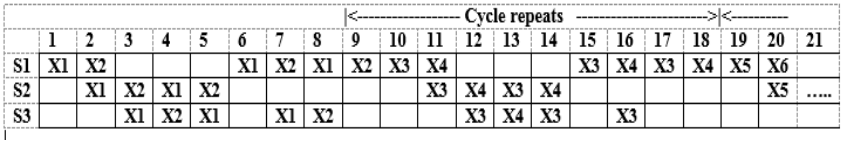
Course Code: 24AM6PCPAP **Max Marks: 100**

Course: Parallel Architectures and Programming

Instructions: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

Important Note: Completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages. Revealing of identification, appeal to evaluator will be treated as malpractice.

		UNIT - I	CO	PO	Marks
1	a)	Elucidate how Flynn's taxonomy categorizes parallel computing architectures based on instruction and data streams.	CO1	PO2	6
	b)	The five stages of the processor have the following latencies: **Instruction** | **Fetch** | **Decode** | **Execute** | **Memory** | **Write back** **A** | 310ps | 420ps | 370ps | 500ps | 200ps **B** | 150ps | 100ps | 200ps | 190ps | 140ps Each stage in the pipeline costs 10ps extra for the registers between stages. Calculate the given below parameters for both Non-pipelined processor and Pipelined processor: i. Cycle time ii. Latency iii. Throughput iv. What is the cycle time, latency and throughput, when one of the longest stage in the pipeline is split into 2 equal halves.	CO1	PO3	8
	c)	An examination of a program has revealed a speedup of 2 when executed on 4 cores. Determine the serial fraction according to Amdahl's and Gustafson's laws.	CO2	PO3	6
		UNIT – II			
2	a)	Explain how optimal load control in a multiprogrammed multiprocessing virtual memory system ensures efficient utilization of CPUs and I/O resources.	CO2	PO2	6
	b)	Explain virtual memory system and three main types of locality of reference with a suitable example.	CO2	PO2	6
	c)	Elaborate the principles of cache mapping techniques and their impact on parallel computing performance.	CO1	PO2	8

			OR			
	3	a)	Differentiate FIFO (First-In-First-Out) and LRU (Least Recently Used) page replacement algorithms with an example.	*CO2*	*PO3*	**7**
		b)	Explain virtual to real page address translation in paged memory system and six fields in page map entry.	*CO1*	*PO2*	**6**
		c)	Explain the design considerations of the input-output interface in a computing system and discuss the contribution of Direct Memory Access (DMA) controllers in optimizing data transfer.	*CO2*	*PO3*	**7**
			UNIT - III			
	4	a)	Explain basic linear pipelining with time space diagram.	*CO2*	*PO2*	**6**
		b)	For the given sequence of stages draw the reservation table and find the forbidden latency. Sa, Sc, Sa, Sb, Sc, Sb	*CO2*	*PO2*	**4**
		c)	Consider the following pipeline reservation table clock cycles / Stages: 1 2 3 4 5 6 7 8 S1: x . . . . x . x S2: . x . x . . . . S3: . . x . x . x . (a) What are the forbidden latencies and permissible latencies? (b) Draw the state transition diagram. (c) List simple cycles and greedy cycles. (d) Determine the minimal average latency.	*CO3*	*PO4*	**10**
			OR			
	5	a)	Elaborate pipeline classification schemes with suitable examples.	*CO3*	*PO3*	**6**
		b)	Draw a 4-stage floating point adder for the given example and explain the stages involved. A= $0.912 * 10^2$ B= $0.8523 * 10^3$	*CO3*	*PO3*	**10**
		c)	Find the latency cycle and average latency for the given reservation table.  	*CO3*	*PO3*	**4**
			UNIT – IV			
	6	a)	How does parallelism differ in GPUs and CPUs. Identify the key factors that contribute to this distinction.	*CO3*	*PO5*	**6**
		b)	Elucidate the memory hierarchy along with its characteristics to address the diverse requirements of GPU workloads.	*CO3*	*PO5*	**6**

		c)	With a neat sketch explain the concepts of threads, grids, and blocks in CUDA. Illustrate how computations are divided into threads to facilitate parallel processing.	*CO3*	*PO5*	**8**
			UNIT – V			
	7	a)	What are the key architectural components of CUDA, and how do they contribute to the parallel processing capabilities of GPUs?	*CO3*	*PO5*	**6**
		b)	Write and explain a CUDA program to add two matrices considering multi-dimensional blockIdx, threadIdx and blockDim.	*CO3*	*PO5*	**7**
		c)	Explain the steps to execute a CUDA program with a neat diagram.	*CO3*	*PO5*	**7**

****