VIT
Vellore Institute of Technology
CHENNAI

Continuous Assessment Test (CAT) – I - JAN 2025

Programme	:	B.Tech CSE	Semester	:	Winter Sem. 24-25
Course Code & Course Title	:	BCSE414L & High Performance Computing	Class Number	:	CH20242505011992
Faculty	:	Dr. Manjula V	Slot	:	A2+TA2
Duration	:	90 minutes	Max. Mark		50

Answer all questions

Q. No	Sub Sec.	Description	Marks
1		Assume two matrixes of size 256x256. Explain how to perform basic multiplication operations on these matrixes using Flynn's taxonomy with appropriate illustration.	10
2	a	Consider quad core processor, and each processor has its own cache and shares common memory. Illustrate the need for cache coherency for the given operations using appropriate protocol of your choice with proper justification. [4M] P1: Write 1050 to A1 Memory location P1: Read A1 P4: Read A1 P4: add 100 with A1 and Write its result to A1 P4: Write 1150 to A2	4
	b	The analysis of binary search program has shown a speedup of 4 when running on 8 cores. (i) Identify the serial fraction according to Gustafson's law with appropriate mathematical formulae. [3M] (ii) Now assume that the parallel work per processor is fixed. Compute the speed-up when using 8 and 16 processors according to Gustafson's law. [3M]	6
3	a	A research institution is planning to deploy a High-Performance Computing (HPC) system for genomic analysis, which requires processing massive datasets efficiently. To ensure optimal performance, the team must consider the critical properties of HPC architecture and operational factors. Identify and describe three key properties that determine the performance of an HPC system. Provide a formula showing how these properties relate to overall system performance. The institution faces challenges in maintaining consistent performance due to environmental and workload-related factors. Name and explain two additional factors that can influence the operation of the HPC system in this scenario. [5M]	10

	b.	A supercomputer has 2,000 nodes, and each node is equipped with 2 processors. Each processor can perform 6 GFLOPS under ideal conditions.
		  i. What is the total peak performance of the supercomputer in PFLOPS (petaflops)? [2M]
		  ii. If the supercomputer uses 64 cores per node, how does that change the peak performance?[1M]	

	c.	Discriminate Strong Scaling Versus Peak Scaling. [2M]	
4	a)	List the primary components of the modern rendering pipeline with GPU. [2M]	10
	b)	Discuss the process and any variants of the process in detail to integrate multiple GPU with one CPU in a single node for improving acceleration with simplified modules. [4M]	
	c)	Explain the necessity of TPUs and discuss the differences with GPUs in terms of working methodology. [4M]	
5	a)	A financial firm is developing a real-time trading platform where high-frequency trades need to be processed rapidly. The system will handle large amounts of market data, execute transactions, and update positions in real-time. The firm decides to use a Symmetric Multiprocessing (SMP) system for its processing needs. The platform is required to be highly responsive and operate with low latency. Why might an SMP system be well-suited for this real-time trading application? [3M]	10
	b)	List the Challenges and potential limitations of SMP in terms of scalability for this application. [4M]	
	c)	Describe how does the shared memory architecture of SMP impact the performance of concurrent tasks in this trading platform with neat sketch.[3M]	

**********All the best *************