GUJARAT TECHNOLOGICAL UNIVERSITY

Program Name: Master of Engineering

Level: PG

Branch: Information Technology

Course / Subject Code : ME01000741

Course / Subject Name : High Performance Computing and Cloud Computing

w. e. f. Academic Year:	A.Y. 2024-25
Semester:	1st Semester
Category of the Course:	PEC

Prerequisite:	**Concept of Parallel Processing and Operating System**
Rationale:	High-performance computing (HPC) involves the specialized use of supercomputers and computer clusters, utilizing parallel programming to accelerate computations. This course covers cutting-edge technologies and programming skills essential for HPC.

Course Outcome:

After Completion of the Course, Student will able to:

No	Course Outcomes	RBT Level
1	Understand the fundamental concepts and new trends of high-performance computing and Cloud computing models and architectures.	U
2	Design, implement, and analyze with OpenMP and MPI for shared-memory system and Distributed Memory	A
3	Learn cutting-edge cloud computing techniques using Amazon AWS and OpenSource tool .	A
4	Perform task using Hadooop cluster environment	E

**Revised Bloom's Taxonomy (RBT)*

Teaching and Examination Scheme:

Teaching Scheme (in Hours)			Total Credits L+T+ (PR/2)	Assessment Pattern and Marks				Total Marks
				Theory		Tutorial / Practical		
L	T	PR	C	ESE (E)	PA / CA (M)	PA/CA (I)	ESE (V)	
3	0	2	4	70	30	20	30	150

Course Content:

Unit No.	Content	No. of Hours	% of Weightage
1.	**Introduction to High Performance Computing** Definition of High Performance Computing, History and latest developments in HPC, Moore's law, Introduction to scientific computing, Challenges with setting up HPC Data centres including storage, power supply and thermal management, topology of processors.	3	10
2.	**Cluster Computing** Clustering Models, Cluster Computing Architectures and key factors, types of clusters, Mission critical Vs Business Critical Applications. Cluster Middleware and SSI, Need of Resource Management and Scheduling.Fault Detection and Masking Algorithms, Check pointing, Heartbeats, Watchdog Timers, Fault recovery through Failover and Failback Concepts.	5	15
3	Introduction to Hadoop Hadoop Distributed File System Hadoop Architecture,MapReduce and HDFS Introduction to Pig,Hive,Hbase	5	10
4	Understanding the MapReduce Internal Components Hbase MapReduce and Hive Overview, Pig Overview,Sqoop Overview Flume Overview Real Time Example in Hadoop Apache Log viewer Analysis Market Basket Algorithms Installing Hadoop Eco System and Integrate With Hadoop Overview of Mahaut,Horton Hadoop Commands usage Monitoring Hadoop Cluster with Ganglia,Nagios,JMX Hadoop Configuration management Tool Hadoop Benchmarking	7	15
5.	**Introduction to Cloud Computing:** •Defining Clouds • Cloud Providers • Consuming Cloud Services • Cloud Models – Infrastructure as a service, Platform as a service, Software as a service • Hypervisor	5	10

	• CPU Virtualization and advantage of Virtual Machine •Containerization		
6.	**Nature of Cloud:** •Tradition Data Center • Cost of Cloud Data Center • Scaling computer systems • Cloud work load • Managing data on clouds • Public, private and hybrid clouds	5	12
7.	• Administering cloud services Identity and Access Management Trust Boundaries and IAM, IAM Challenges, Relevant IAM Standards and Protocols for Cloud Services, IAM Practices in the Cloud, Cloud Authorization Management	5	13
8.	Parallel computing using OpenMP Introduction, fork and join, synchronizations, race conditions, compiler directives, heap and stack memory, barriers, overheads, reductions, data dependencies, thread-safe functions, pseudo-random number generators in parallel, example codes, and demonstrations. Parallel computing using MPI Introduction, message passing, domain decomposition, MPI communicators, MPI modules and functions (broadcast, reduce, allreduce, MPI send and receive), master– worker paradigm	5	15

Suggested Specification Table with Marks (Theory):

Distribution of Theory Marks (in %)					
R Level	**U Level**	**A Level**	**N Level**	**E Level**	**C Level**
10	**30**	**30**	**20**	**10**	**---**

Where R: Remember; U: Understanding; A: Application, N: Analyze and E: Evaluate C: Create (as per Revised Bloom's Taxonomy)

References/Suggested Learning Resources:

(a) Books:

1. Rajkumar, High Performance Cluster Computing: Architectures and Systems, Vol. 1 Pearson Education
2. Georg Hager and Gerhard Wellein, Introduction to High Performance Computing for Scientists and Engineers, CRC Press

3. Cloud Computing Explained: Implementation Handbook for Enterprises, John Rhoton, Publication Date: November 2, 2009
4. M. Quinn, Parallel Programming in C with MPI and OpenMP, McGraw-Hill Education, 2003.
5. G. Karniadakis, R. Kirby II, Parallel Scientific Computing in C++ and MPI, Cambridge University Press, 2003.
6. The Art of Multi-Processor Programming, 2nd ed. Morgan-Kaufmann, 2021.

(b) Open source software and website:

- NPTEL Course : High Performance Computing

1. Course Name: High Performance Computing
2. Link: https://nptel.ac.in/courses/106/108/106108055/

3. Course Name: High Performance Computing Architecture
4. Link: https://nptel.ac.in/courses/106/105/106105033/

Web resource

- https://www.open-mpi.org/
- https://www.mpich.org/
- https://research.cs.wisc.edu/htcondor/
- https://en.wikipedia.org/wiki/Cilk
- https://github.com/ioannischristou/popt4jlib

Suggested Course Practical List: (List can be change according to Latest Development)

1. 1 To study the basic commands of linux.
2. 2 To establish Beowulf Cluster using MPI(Message Passing Interface) Library.
3. 3 Installation and configuration of Alchemi Grid.
4. 4 Running a sample application on Alchemi Grid and analysing it.
5. 5 To study a Grid Simulation Toolkit.
6. 6 To run two sample programs using GridSim Toolkit.
7. Reading and Writing the files in HDFS using java program
8. Write a Sample Hadoop Examples for Word count program and Population problem.
9. To study a Cloud Simulation Toolkit.

Candidate need to setup following as a part of Laboratory work:

- Apache Hadoop Installation
- Zookeeper Installation
- Hbase Installation
- Hive Installation

- Pig Installation
- Sqoop Installation
- Installing Mahout
- Import the data in HDFS
- Setup Private cloud using opensource tool.

List of Laboratory/Learning Resources Required: Programming Languages – Python / JAVA etc.

* * * * * * *