Pre-Requisite(s): 20CYS203 - Operating System, 20CYS204 - Database Management System

Course Objectives

- Introduction to distributed systems and cloud computing.

- Understand different cloud architectures and technology.

- Illustrate the use of Hadoop clusters and Peer to Peer Systems

Course Outcome

CO1: Classify and describe the architecture and taxonomy of parallel and distributed computing, including shared and distributed memory, and data and task parallel computing

CO2: Characterize the distinctions between Infrastructure, Platform and Software as a Service (IaaS, PaaS, SaaS) abstractions, and Public and Private Clouds, and analyze their advantages and disadvantages.

CO3: Exploring Hadoop clusters and Peer to Peer Systems

CO-PO Mapping

PO/PSO CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2
CO1	2	1			1	1							2	2
CO2	2	2	1		1	1							2	2
CO3	2	1	2		1	1							2	2

Syllabus

Introduction to distributed systems, Distributed computing paradigms, Inter process communication mechanisms, Process models in distributed systems, The CAP theorem, Consistency models and Replication, Consensus algorithm: Clock Synchronization – Logical clocks – Mutual Exclusion, global positioning of nodes, Distributed Commit protocols – 2PC, 3PC, Check-pointing and Recovery, Election algorithms, Failure Models, Paxos algorithm-Apache Zookeeper, Distributed file system – Eg: CODA and Ceph, Distributed storage implementation – Data sharding, nosql key value stores and its properties – Eg: Google Big Table, Amazon DynamoDB. Cloud computing benefits and its challenges, Types – Private, Public and Hybrid clouds, Models – IaaS, PaaS and SaaS. Role of virtualization in enabling

the cloud computing; Business Agility: Benefits and challenges to cloud architecture. AWS cloud services and management – scalability, availability, concurrency with practical aspects, REST API services including load balancing, server authentication and debug handling, AWS Zelkova for Provable Security. Hadoop cloud computing framework – HDFS and MapReduce, Cloud data processing using Pig and Hive, Amazon EMR for creating Hadoop clusters within AWS. Peer to Peer Systems – Napster, Gnutella, FastTrack, BitTorrent, Distributed Hash Tables, IPFS.

Text Book(s)

1. *Andrew S. Tannenbaum and Maarten van Steen, Distributed Systems: Principles and Paradigms, Third Edition, Prentice Hall, 2017.*

2. *Ronald L. Krutz, Russell Dean Vines. Cloud Security: A comprehensive Guide to Secure Cloud Computing, Wiley India 2010.*

Reference(s)

1. *Ajay D. Kshemkalyani and Mukesh Singhal, Distributed Computing: Principles, Algorithms, and Systems, Cambridge University Press, 2011.*

2. *Garg VK, Garg VK. Elements of distributed computing. John Wiley & Sons; 2002.*

3. *George Coulouris, Jean Dollimore, Tim Kindberg and Gordon Blair, Distributed Systems: Concepts and Design, Fifth Edition, Pearson Education, 2017.*

4. *Fokkink W. Distributed algorithms: an intuitive approach. Second Edition, MIT Press; 2018.*

Evaluation Pattern

Assessment	Internal	External
Periodical 1	10	
Periodical 2	10	
Continuous Assessment (Theory) (CAT)	15	
Continuous Assessment (Lab) (CAL)	30	
End Semester		35

*CAT – Can be Quizzes, Assignment, Projects, and Reports.