Nirma University
Institute of Technology, School of Technology
MTech Computer Science and Engineering / MTech Computer Science and Engineering (Information and Network Security)

Semester – II

L	T	P	C
3	0	2	4

Course Code	3CS12D105
Course Title	Distributed Systems

Course Learning Outcomes (CLOs):

At the end of the course, students will be able to

1. comprehend the computing models for distributed environment
2. analyse the distributed systems in the context of various performance parameters
3. apply the modern concepts of distributed databases and distributed file systems

Syllabus:

Teaching Hours

Unit I

Introduction and Architectures: Features and types of distributed systems, Distributed Models of Computation, Architectures: Regular graphs, random graphs, power-law, and small-world networks, middleware, and self-management in distributed systems, causality & logical time — **5**

Unit II — **7**

Processes, Communications and Programming models for Computing: Multithreaded clients and services, virtualization and virtual machines, and code migration, lightweight RPC, message and stream-oriented communication and multicast communication, Data-Intensive Computing, MapReduce: Core design elements, In-memory cluster compute with Spark

Unit III

Logical Time, Global State & Snapshot and Distributed Mutual Exclusion: — **6**

Virtual Time, Physical Clock Synchronization, Global State and Snapshot Recording Algorithms, Distributed Mutual Exclusion with Token and Non-Token based Approaches, Leader Election in Rings, Consensus and Agreement Algorithms, Deadlock detection and recovery in Distributed Systems

Unit IV

Consistency and Replication: Consistency models: strong and weak, Scalable Causal Consistency, Highly reliable distributed coordination with Zookeeper, Replication management, Distributed Replication, PAXOS and RAFT Algorithms — **6**

Unit V

Transactional Systems and Distributed Databases: Distributed Transactions, Atomic Commit and Optimistic Concurrency Control: FaRM, Spark: Resilient distributed dataset (RDD), Operations: transformations & actions, Spark Streaming, Spanner: Google's Globally-Distributed Database, SNOW Theorem **6**

Unit VI **5**

Fault Tolerance and Security: Failure models, failure detection, algorithms for fault tolerance, and recovery from failure in distributed systems, Authentication in Distributed Systems, Distribution of security mechanisms, access control, and security management

Unit VII **10**

File Systems and Distributed Storage Systems: Network File System, Andrew File system, Google File System, Hadoop Distributed File System

Case Studies: Internet content delivery systems, Distributed ledgers and Block chains, Distributed machine learning, Scale-out key-value storage: Dynamo

Self-Study:

The self-study contents will be declared at the commencement of semester. Around 10% of the questions will be asked from self-study contents.

Laboratory Work:

Laboratory work will be based on above syllabus with minimum 5 experiments to be incorporated.

Suggested Readings^:

1. Andrew S. Tanenbaum and Maarten van Steen, Distributed Systems: Principles and Paradigms, Createspace
2. George Coulouris, Jean Dollimore, Tim Kindberg, Gordon Blair, Distributed Systems: Concepts and Design, Addison Wesley
3. Ajay D. Kshemkalyani and Mukesh Singhal, Distributed Computing: Principles, Algorithms, and Systems, Cambridge University Press
4. Kenneth P Birman, Guide to Reliable Distributed Systems: Building High-Assurance Applications and Cloud-Hosted Services, Springer
5. Relevant research papers in the area of distributed systems

L=Lecture, T=Tutorial, P=Practical, C=Credit

^this is not an exhaustive list