GUJARAT TECHNOLOGICAL UNIVERSITY

COMPUTER ENGINEERING (SOFTWARE ENGINEERING) (02)
DISTRIBUTED DATABASE APPLICATION SYSTEM
SUBJECT CODE: 2740202
M.E. 4TH SEMESTER

Type of course: Elective

Prerequisite: Knowledge of DBMS, Distributed Systems.

Rationale: Distributed Systems are gaining popularity due to various advantages it offers. Database is also getting distributed. When database is distributed, the concepts of database need to be revisited, the student should be made aware of the concepts such as architecture, how to distribute database, database control, optimizing query, controlling replication, handling concurrency and deadlock.

Teaching and Examination Scheme:

Teaching Scheme			Credits	Examination Marks						Total Marks
				Theory Marks		Practical Marks				
L	T	P	C	ESE (E)	PA (M)	ESE (V)		PA (I)		
						ESE	OEP	PA	RP	
3	2#	0	4	70	30	30	0	10	10	150

Content:

Sr. No.	Content	Total Hrs	% Weightage
1	**Introduction:** Distributed Data Processing, Promises of DDBs, Complicating Factors, and Problem Areas.	2	5
2	**Distributed DBMS Architecture:** DBMS Standardization, Architectural Models for Distributed DBMSs, Distributed DBMS Architecture.	2	5
3	**Distributed Database Design:** Alternative Design Strategies, Distribution Design Issues, Fragmentation, Distribution Transparency, Impact of distribution on user queries, and Allocation.	4	10
4	**Semantic Data Control:** Authentication and Access rights, View Management, Data Security, Semantic Integrity control & its enforcement	2	5
5	**Overview of Query Processing:** Query Processing Problem, Objective of Query Processing, Complexity of Relational Algebra Operations, Query processing in centralized system, Query processing in distributed system, Characterization of Query Processors, Layers of Query Processing.	4	10
6	**Query Decomposition & Data Localization:** Query Decomposition, Localization of Distributed Data.	2	5
7	**Optimization of Distributed queries:** Query optimization, Centralized Query optimization, Join Ordering in Fragmented Queries, Distributed Query Optimization algorithms.	3	5
8	**Introduction to Transaction Management:** Definition Of Transaction, Properties of Transactions, Types of Transactions.	2	5

9	**Distributed Concurrency Control:** Serializability Theory, Taxonomy of Concurrency control Mechanisms, Lock based Concurrency control Mechanisms, TimeStamp-Based Concurrency control Algorithms, Optimistic Concurrency control Algorithms, Deadlock Management, Relaxed Concurrency Control.	4	10
10	**Distributed DBMS Reliability:** Reliability Concepts & Measures, Failures & Fault Tolerance in Distributed systems, Failures in Distributed DBMS, Local Reliability Protocols, Distributed Reliability Protocols, Dealing with site failures, Network Partitioning.	4	10
11	**Parallel Database Systems**: Database Servers, Parallel Architectures, Parallel DBMS Techniques, Parallel Execution problems, Parallel Execution For Hierarchical Architecture.	3	5
12	**Distributed Object Database Management Systems:** Fundamental objects concepts & object models, Object distribution design, Architectural issues, Object Management, Distributed Object Storage, Object Query Processing, Transaction Management.	3	5
13	**Database Interoperability:** Database Integration, Query Processing, Transaction Management, Object Orientation & Interoperability.	2	5
14	**Implementation:** Alternatives , Overview of Java Messaging Service (JMS), J2EE support for Distributed Process platform requirements, Microsoft .Net support for Distributed Process platform requirements	3	10
15	**Current Issues:** Data Delivery Alternatives, Data Warehousing, World Wide Web, Push-based Technologies, Mobile Databases.	2	5

Reference Books:

1. Principles of Distributed Database Systems by M. Tamer Ozsu, Patrick Valduriez (Pearson Publication)
2. Distributed Database Management Systems- A Practical Approach by Saeed K Rahimi, Frank S Haug (Wiley Publication)
3. Distributed Databases Principles and Systems by Stefano Ceri, Giuseppe Pelagatti (Mcgraw Hill Publication)

Course Outcome:
After learning the course the students should be able to:

1. Understand Distributed database systems (DDBMS)
2. Architecture & Design of DDBMS
3. Query Processing & Optimization
4. Concurrency control & reliability issues in DDBMS
5. Use tools for implementing DDBMS applications.

List of Tutorials:
Tutorials/Problems based on content such as

- fragmentation,
- database design ,
- query processing and decomposition,
- Serializability.
- Deadlock etc

Implementation such as

- connection of two different database on same machine,
- connection of two different databases on different machines,
- Transactions,
- creating views,
- materialized view,
- implementing small application of DDBMS using tools such as J2EE, .NET etc

Major Equipment: Networked Lab, RDBMS software, Open Source Software

List of Open Source Software/learning website:

http://infolab.stanford.edu/cs347.2001.spring/lectures.html
http://computingnow.computer.org/web/computingnow
http://www.csee.umbc.edu/portal/help/oracle8/server.815/a67784/ds_ch5.htm

Review Presentation (RP): The concerned faculty member shall provide the list of peer reviewed Journals and Tier-I and Tier-II Conferences relating to the subject (or relating to the area of thesis for seminar) to the students in the beginning of the semester. The same list will be uploaded on GTU website during the first two weeks of the start of the semester. Every student or a group of students shall critically study 2 papers, integrate the details and make presentation in the last two weeks of the semester. The GTU marks entry portal will allow entry of marks only after uploading of the best 3 presentations. A unique id number will be generated only after uploading the presentations. Thereafter the entry of marks will be allowed. The best 3 presentations of each college will be uploaded on GTU website.