GUJARAT TECHNOLOGICAL UNIVERSITY

Master of Engineering
Subject Code: 3722321
Semester II
SUBJECT NAME: Distributed Databases

Type of course: Elective

Prerequisite: Database Management System, Computer Network, 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)	
3	0	2	4	70	30	30	20	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 DBMS, Distributed DBMS Architecture	2	5%
3	**Distributed Database Design:** Alternative Design Strategies, Distribution Design Issues, Fragmentation, Distribution Transparency, and Allocation.	5	10%
4	**Semantic Data Control:** Authentication and Access Rights, View Management, Data Security, Semantic Integrity Control & its Enforcement	2	4%
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**	**4%**
7	**Optimization of Distributed Queries:** Query Optimization, Centralized Query Optimization, Join Ordering in Fragmented Queries, Distributed Query Optimization Algorithms.	**5**	**10%**
8	**Introduction to Transaction Management:** Definition of Transaction, Properties of Transactions, Types of Transactions.	**2**	**4%**
9	**Distributed Concurrency Control:** Serializability Theory, Taxonomy of Concurrency Control Mechanisms, Lock based Concurrency Control Mechanisms, Time Stamp based Concurrency Control Algorithms, Optimistic Concurrency Control Algorithms, Deadlock Management, Relaxed Concurrency Control	**5**	**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.	**5**	**10%**
11	**Parallel Database Systems:** Database Servers, Parallel Architectures, Parallel DBMS Techniques, Parallel Execution Problems, Parallel Execution for Hierarchical Architecture.	**5**	**10%**
12	**Implementation:** Overview of Java Messaging Service(JMS), J2EE Support for Distributed Process Platform Requirements, Microsoft .Net Support for Distributed Process Platform Requirements, RMI, CORBA	**6**	**12%**
13	**Current Issues:** Data Delivery Alternatives, Push-based Technologies, Mobile Databases, Models and Current Trends in Improving the Performance of Distributed Database.	**3**	**6%**

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:

Sr. No.	CO statement	Marks % weightage
CO-1	Understand Distributed database systems (DDBMS), its architecture, design & security issues.	25

CO-2	Apply the concepts of query processing and decomposition to different queries.	20
CO-3	Understand the issues related to DDBMS such as reliability, concurrency control and related algorithms.	25
CO-4	Understand parallel database system and its architecture and latest trends in subject.	15
CO-5	Implement various DDBMS problems using JMS, J2EE, RMI, CORBA	15

Suggested Practical List:

1. Write an application using JAVA/.NET which access two different databases created using same DBMS and implements DDL and DML commands.
2. Write an application using JAVA/.NET which access two different databases created using two different DBMS and implements DDL and DML commands.
3. Write an application using JAVA/.NET which shows implementation of horizontal fragmentation.
4. Write an application using JAVA/.NET which shows implementation of vertical fragmentation.
5. Write an application using JAVA/.NET which shows implementation of hybrid fragmentation.
6. Write an application using JAVA/.NET to demonstrate distributed query processing.
7. Write an application using JAVA/.NET to demonstrate ACID properties implementation of transaction.
8. Write an application using JAVA/.NET to demonstrate concurrency control mechanism.
9. Write an application using JAVA/.NET to demonstrate two phase commit protocol.
10. Write an application using JAVA/.NET to demonstrate three phase commit protocol.
11. Write an application using CORBA to demonstrate distributed database system.
12. Write an application using RMI to demonstrate distributed database system.

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