GUJARAT TECHNOLOGICAL UNIVERSITY

INFORMATION TECHNOLOGY (23)
ADVANCE DATABASE
SUBJECT CODE: 2722311
SEMESTER: II

Type of course: Master of Engineering

Prerequisite: Database Management System

Rationale:
At the graduate level, student had studied database management system. By looking into the fast growing development in the field of information technology, advance database in different applications and query language are used. This course covers salient features of various types of databases, system architecture of various types of databases, of transaction management, data warehousing.

Teaching and Examination Scheme:

Teaching Scheme			Credits	Examination Marks						Total Marks
L	T	P	C	Theory Marks		Practical Marks				
				ESE (E)	PA (M)	ESE (V)		PA (I)		
						ESE	OEP	PA	RP	
3	2#	0	4	70	30	20	10	10	10	150

Content:

Sr. No.	Content	Total Hrs	% Weightage
1	Database System Architectures: Centralized Systems, Client-server System Architecture, Parallel Systems, Distributed Systems, Network Types	2	10
2	Parallel Databases: I/O Parallelism, Inter-query and Intra-query Parallelism, Inter-operation and Intra-operation parallelism.	2	10
3	Distributed Databases: Introduction - Distributed Data Processing, Complicating Factors, Problem Areas. Distributed DBMS architecture. Distributed Database Design – Design Strategies, Design Issues, Fragmentation, Allocation.	10	20
4	Data Warehouse and OLAP: Introduction- Operational Database Systems and Data Warehouses. Multidimensional Data Model - Star, Snowflake, and Fact Constellation Schemas. Architecture, Indexing OLAP and Query Processing, Introduction to Analytical Mining.	8	20
5	Object-Based Databases and XML: Need for complex data types, Persistent Programming Languages, Object-Oriented versus Object-Relational. XML – Structure of XML Data, Document Schema, Application Program Interface, Storage of XML Data, Applications	10	20
6	Advanced Data Types and New Applications: Time in Database, Spatial and Geographic Data, Multimedia Database,	3	10

	Mobility and Personal Database		
7	Advanced Transaction Processing: Transaction Processing Monitors, Transactional Workflow, Main memory database, Real time transaction systems, Long duration transactions, Transaction management in multi databases.	5	10

Reference Books:

1. Database System Concepts, Abraham Silberschatz, Henry F. Korth & S. Sudarshan, McGraw Hill.

2. Principles of Distributed database systems, By M. tamer Ozsu, Petrick Valduriez, Pearson

3. Distributed Database Systems By Chhanda Ray, Pearson.

4. Jiawei Han and Micheline Kamber, "Data Mining – Concepts and Techniques", Second Edition, Elsevier.

Course Outcome:

Upon the successful completion of this course, students should be able to:

1. Explain various types of database system architecture.
2. Design queries against a distributed database management system.
3. Perform queries against database designed with object-relational extensions.
4. Develop professional attitude towards the development of database application.
5. Develop and query XML files.
6. Design multi-dimensional data model.

List of Experiments:

1. Assume we have a global conceptual schema that contains the following table with the key underlined: Employee(Eno,Ename,Title,Dno). Also assume that we horizontally fragment the table as follows:
   Employee1(Eno;Ename; Title;Dno), where 1<= Dno <=10
   Employee2(Eno;Ename; Title;Dno), where 11 <= Dno <=20
   Employee3(Eno;Ename; Title;Dno), where 21 <= Dno <=30
   In addition, assume we have 4 sites that contain the following fragments:
   Site1 has Employee1
   Site2 has Employee2
   Site3 has Employee2 and Employee3
   Site4 has Employee1

2. Implement at least 5 suitable queries using suitable database system on Employee fragment.
3. We are given the following three relations with their keys underlined:
   Supplier(Sno,Sname,City,State)
   Part(Pno,Pname,Color)
   Supplier-Part(Sno,Pno,Qty).

   We know that Suppliers can supply many Parts and many Suppliers can supply a Part. Assume the Supplier table is horizontally fragmented using the predicates: State =Maharashtra and State = Gujarat. We can also assume that Suppliers are evenly located in only those two states.

In addition, the Part table is horizontally fragmented using the predicates: 1<= Pno<=100,101<=Pno <=200, 201<=Pno<=300, 301<=Pno<=400, 401<=Pno<=500. Part numbers are continuous from 1 to 500, inclusive.

Now we are to horizontally fragment the Supplier- Part relation according to your choice.

4. Implement at least 5 suitable queries using suitable database system for practical #3.
5. To design data cube by identifying dimensions and measures for star schema.
6. To design data cube by identifying dimensions and measures for snowflake schema.
7. Case Study: Applications of Data Mining in field of finance, security, retail-market and e-commerce.
8. Create a table employee having dept_id as number datatype and employee_spec as XML datatype (XML Type). The employee_spec is a schema with attributes emp id, name, email, acc_no, managerEmail, dateofjoining. Insert 10 tuples into employee table.
   a. Fire the following queries on XML database.
   b. Retrieve the names of employee.
   c. Retrieve the acc_no of employees.
   d. Retrieve the names, acc_no, email of employees.
   e. Update the 3rd record from the table and display the name of an employee.
   f. Delete 4th record from the table.

Open Ended Problems:

• How to manage heterogeneity in distributed database? For example database fragmented to different site and each site having different database architecture. To execute a global query over number of sites having different database architecture.

• Take appropriate dataset (i.e Finance, Whether, Mall etc.) having more than 25000 data and design data cube by identifying dimensions and measures for star/snowflake schema

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.