UNIT II Query processing and decomposition

Query Processing Objectives

- – Minimize response time
- – Reduce communication cost
- – Efficient resource utilization
- – Improve throughput and scalability
- – Maintain correctness and consistency

Characterization of Query Processors

- – Query parsing
- – Query validation
- – Query transformation
- – Query optimization
- – Query execution

Layers of Query Processing

- 1. User Interface Layer
- 2. Query Decomposition Layer
- 3. Data Localization Layer
- 4. Global Optimization Layer
- 5. Local Optimization & Execution Layer

Query Decomposition

- – Parsing and semantic analysis
- – Query normalization
- – Query rewriting
- – Generation of relational algebra expressions

Localization of Distributed Data

- – Mapping global queries to local fragments
- – Fragment identification
- – Data allocation awareness
- – Reconstruction of distributed relations

Query Optimization

- – Selection of efficient execution plans
- – Cost estimation
- – Join ordering
- – Access path selection

Centralized Query Optimization

- – Single-site database environment
- – Cost-based optimization
- – Heuristic optimization
- – Join optimization techniques

Distributed Query Optimization

- – Multiple-site execution planning
- – Communication cost consideration
- – Data transfer minimization
- – Parallel execution strategies

Distributed Query Optimization Algorithms

- – Dynamic Programming
- – Greedy Algorithms
- – Branch and Bound
- – Heuristic-based Optimization
- – Semi-Join Strategies

Advantages & Challenges

- Advantages:
- – Better performance
- – Scalability

- Challenges:
- – Network overhead
- – Complex cost estimation
- – Data distribution management

Conclusion

- Query decomposition and optimization are essential for efficient distributed database query execution.