Introduction

A Brief Overview: Parallel Databases and Grid Databases

Parallel Query Processing: Motivations

Parallel Query Processing: Objectives

Forms of Parallelism

Parallel Database Architectures

Grid Database Architecture

Structure of this Book

Summary

Bibliographical Notes

Exercises

Analytical Models

Cost Models

Cost Notations

Skew Model

Basic Operations in Parallel Databases

Summary

Bibliographical Notes

Exercises

Basic Query Parallelism

Parallel Search

Search Queries

Data Partitioning

Search Algorithms

Summary

Bibliographical Notes

Exercises

Parallel Sort and Group-By

Sorting, Duplicate Removal, and Aggregate Queries

Serial External Sorting Method

Algorithms for Parallel External Sort

Parallel Algorithms for GroupBy Queries

Cost Models for Parallel Sort

Cost Models for Parallel GroupBy

Summary

Bibliographical Notes

Exercises

Parallel Join

Join Operations

Serial Join Algorithms
Serial vs. Parallel Execution Scheduling
Scheduling Rules
Cluster Query Processing Model
Dynamic Cluster Query Optimization
Other Approaches of Dynamic Query Optimization
Summary
Bibliographical Notes
Exercises
Grid Databases
Transactions in Distributed and Grid Databases
Grid Database Challenges
Distributed Database Systems and Multidatabase Systems
Basic Definitions on Transaction Management
ACID Properties of Transactions
Transaction Management in Various Database Systems
Requirements in Grid Database Systems
Concurrency Control Protocols
Atomic Commit Protocols
Replica Synchronization Protocols
Summary
Bibliographical Notes
Exercises
Grid Concurrency Control
A Grid Database Environment
An Example
Grid Concurrency Control
Correctness of GCC Protocol
Features of GCC Protocol
Summary
Bibliographical Notes
Exercises
Grid Transaction Atomicity and Durability
Motivation
Grid Atomic Commit Protocol (Grid-ACP)
Handling Failure of Sites with Grid-ACP
Summary
Bibliographical Notes
Exercises
Replica Management in Grids
Motivation
Replica Architecture
Grid Replica Access Protocol (GRAP)
Handling Multiple Partitioning
Summary
Bibliographical Notes
Exercises
Grid Atomic Commitment in Replicated Data
Motivation
Modified Grid Atomic Commitment Protocol
Transaction Properties in Replicated Environment
Summary
Bibliographical Notes
Exercises
Other Data Intensive Applications
Parallel Online Analytic Processing (OLAP) and Business Intelligence
Parallel Multidimensional Analysis
Parallelization of ROLLUP Queries
Parallelization of CUBE Queries
Parallelization of Top-N and Ranking Queries
Parallelization of CUME_DIST Queries
Parallelization of NTILE and Histogram Queries
Parallelization of Moving Average and Windowing Queries
Summary
Bibliographical Notes
Exercises
Parallel Data Mining - Association Rules and Sequential Patterns
From Databases, Data Warehousing, to Data Mining: A Journey
Data Mining: A Brief Overview
Parallel Association Rules
Parallel Sequential Patterns
Summary
Bibliographical Notes
Exercises
Parallel Clustering and Classification
Clustering and Classification
Parallel Clustering
Parallel Classification
Summary
Bibliographical Notes
Exercises
Table of Contents provided by Blackwell's Book Services and R.R. Bowker. Used with permission.