Distributed Database Systems

Autumn, 2008

Welcome!

Crew

- Lecturer: **Prof. Lizhu Zhou 周立柱**
  - DB Group, Institute of Software, Dept. of CS&T, Tsinghua University
  - dcszlz@tsinghua.edu.cn
- Teaching Assistant: **Ju Fan 范举**
  - If you have any question, call 51537853 or 13811942748 or send an e-mail to:
  - fan-j07@mails.tsinghua.edu.cn

Title

Distributed Database Systems

Teaching hours

48

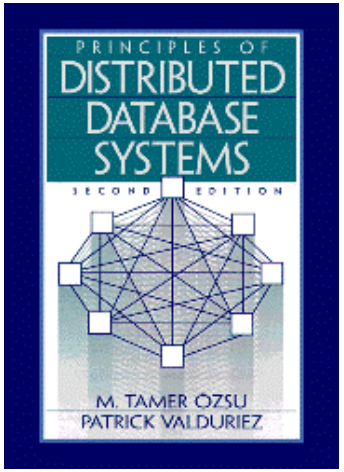
Textbook

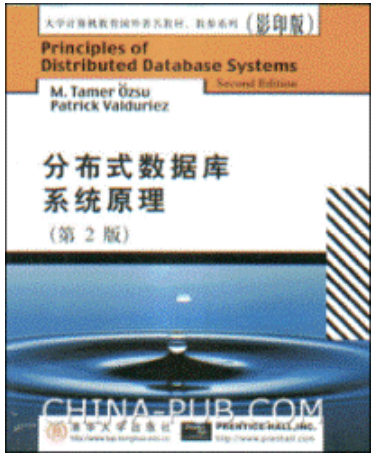
Principles of Distributed Database Systems

(2nd Edition)

M. Tame Özsu
Patrick Valduriez
Prentice-Hall, 1999

Price: RMB 57.00





Reference

Distributed Databases: Principles and Systems

Ceri & Pelagatti McGraw-Hill, 1985

许多中文参考书

Prerequisite

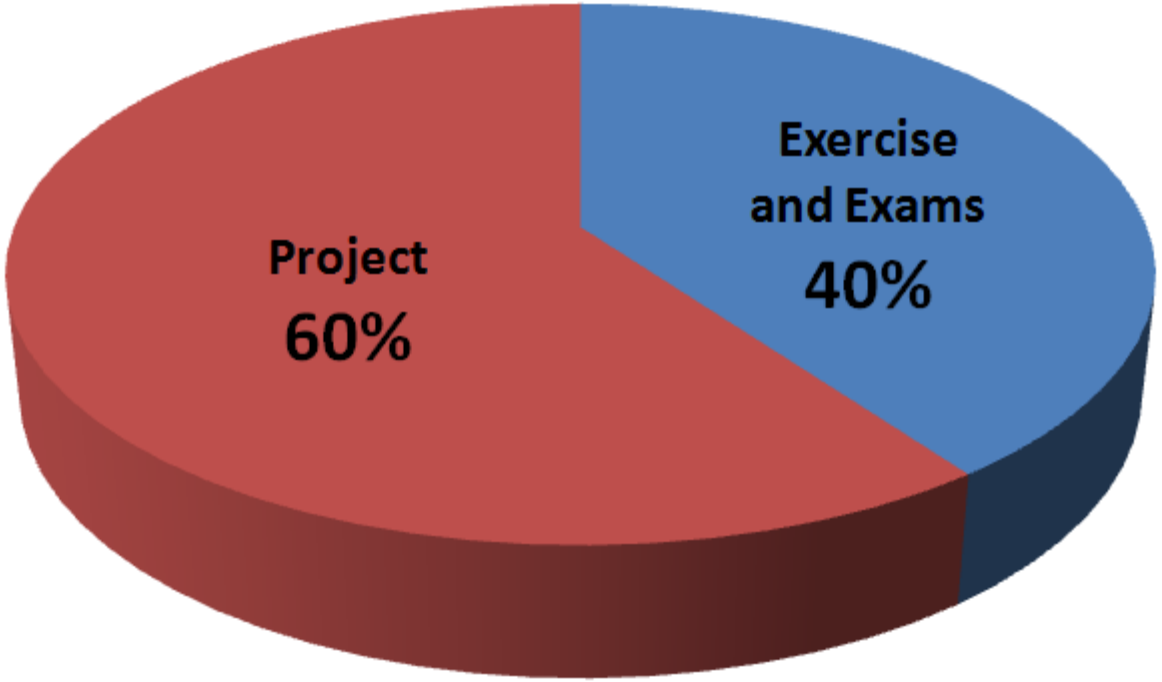
Undergraduate **database** course

Computer networks

Teaching language

English

Evaluation



Contents

Chapter #	Title
1	Introduction
4	Distributed DBMS Architecture
5	Distributed Database Design
7	Overview of Query Processing
8	Query Decomposition and Data Localization
9	Optimization of Distributed Queries
10	Introduction to Transaction Management
11	Distributed Concurrency Control
12	Distributed DBMS Reliability

Course home page

- http://dbgroup.cs.tsinghua.edu.cn/ddb/
- On the home page you can
  - Get notifications
  - Download the latest courseware
  - Download other related materials

Friendly Reminding Messages from the Instructor

Reading for learning

- Problem formulation
  - What is the problem
  - Relevant Concept definition
  - Mathematical representation
- Algorithm description & formal presentation
  - What is the method
  - Writing pseudo code
- Proof of your works
  - Theoretical analysis
  - Experimental study

Are You Scared by DDB Course?

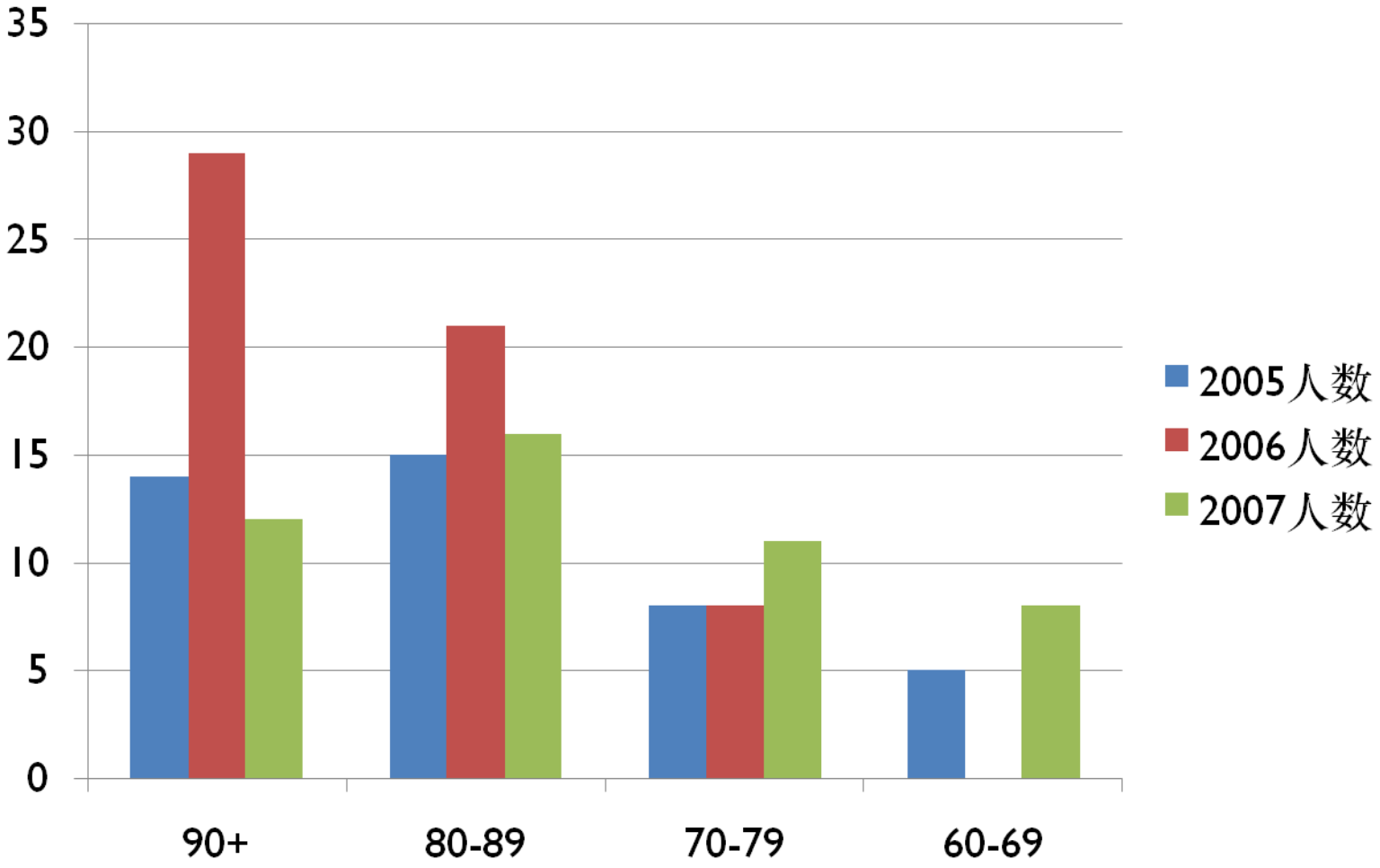
2005（左)-2006(右）成绩分布

- 选课人数：43人
- 平均成绩：83
- 最高分：98
- 90+的人数：14
- 80-89：15
- 70-79：8
- 60-69：5
- 60-：1

- 选课人数：58人
- 平均成绩：89
- 最高分：98
- 90+的人数：29
- 80-89：21
- 70-79：8
- 60-69：0
- 60-：0

2007成绩分布

- 选课人数：49人
- 平均成绩：79.86
- 最高分：98
- 90+的人数：12
- 80-89：16
- 70-79：11
- 60-69：8
- 60-：2


35
30
25
20
15
10
5
0
90+
80-89
70-79
60-69
2005人数
2006人数
2007人数

课程设计中应注意的问题

1. 选课太多，疲于奔命，无暇顾忌
2. 缺乏总体设计，系统划分不当，随意编程，接口混乱
3. 设计目标很高，实现时想一步到位，指望一次成功
4. 选择开发环境欠考虑
5. 代码管理混乱，难以形成正确版本

6. 实现时不考虑调试需求（例如反复建立表格，生成数据字典）
7. 匆忙上马，反复折腾
8. 各自为战，缺少交流
9. 三人平等，没有核心
10. 分工不均，事后抱怨

Good Luck!

The History of Database Field

- 1960's
  - IBM IMS
  - DBTG network data model
- 1970's
  - E. F. Codd RDB model
  - Pioneering systems: INGRES, System R
- 1980's
  - Booming of RDBs
  - Federated, Multi-databse
  - DDB
  - OODB

- 1990's
  - Merge of RDB and OODB
  - Internet Fame
  - P2P model
- 2000's
  - XML DB
  - Web data processing
  - Semantic Web
  - Data Integration

Status of Commercial DDB

- Server side P2P
- Support
  - Horizontal fragmentation
  - 2PC

Why study DDB

- The fundamental principles for distributed systems
- Scientific training on
  - basic concepts
  - problem formulation
  - identification of key technical issues
  - algorithm presentation
  - system level design and implementation