

Book Review:

Principles of Distributed Database Systems, Third Edition

Tamer Özsu, Patrick Valduriez
Springer – 2011
ISBN 978-1-4419-8833-1

We could consider this book as a *classical* one from many points of view. To start with, the book is the third in a series of editions since the first one more than 20 years ago, thus containing the evolution in the field of distributed databases as well as the basic concepts which are not strongly affected by technology evolution. The thorough review of relational databases as well as a discussion of several aspects of distributed systems (as computer networks) in Chapter 2 makes the book highly self-contained, enhancing reading and understanding underlying principles. The book covers all the problems expected to be encountered in distributed databases:

1. Fragmentation (horizontal as well as vertical fragmentation) and allocation (Chapter 3)
2. Replication (Chapter 13)
3. Database queries
 - 3.1. Overview (Chapter 6)
 - 3.2. Decomposition (Chapter 7)
 - 3.3. Optimization (Chapter 8)
4. Database integration/federation
 - 4.1. Integration design (Chapter 4)
 - 4.2. Query processing (Chapter 9)
5. Data and Access Control (Chapter 5)
6. Transactions related issues
 - 6.1. Introduction to transaction management (Chapter 10)
 - 6.2. Concurrency Control (Chapter 11)
 - 6.3. Commit Protocols (Chapter 12)

The book also covers other interesting issues in the context of distributed databases, and the current development/evolution in the field will determine whether or not those issues are going to be relevant in the future as distributed database topics:

1. Parallel database systems (Chapter 14)
2. Distributed Object Database Management (Chapter 15)
3. Peer-to-Peer Data Management (Chapter 16)
4. Web Data Management (Chapter 17)
5. Current Issues: Streaming Data and Cloud (Chapter 18)

where some of these topics are strongly related to the currently so-called *NoSQL databases*.

Authors have made an outstanding and successful effort in detailing and commenting references of related work in each topic at the end of each chapter. Also, references are used along descriptions/explanations in every chapter so the reader is able to use those references to dig in more detail. In this sense, the book style is solid, almost everything is explained in detail and/or in terms of well-known sources.

Most of the book is based on/focused in relational databases. Some exceptions can be found in the final chapters and in the end of some other chapters. Unfortunately, this is not clearly identified, but

can be understood in a thorough (maybe second) specific reading. Other detail which is hard to identify in some parts of the book (mostly at the beginning of some chapters), is whether the description of some problems and solutions is explained in general terms (i.e. distributed systems) or distributed database systems.

The book has supplementary material provided by Springer:

<http://extras.springer.com/2011/978-1-4419-8833-1>

containing figures as well as full presentation slides. One of the authors, M. Tamer Ösu also provides supplementary material at

<https://cs.uwaterloo.ca/~tozsu/ddbook/>

including a way for asking for the solutions to exercises, expected to be used by those who teach a course using this book. There is a link to an Errata page, but empty by the time of this writing.

Summarizing, a very good book, specifically recommended to be used in specific courses on distributed database systems as well as for those who want to start working on this field.

Prof. Fernando G. Tinetti (fernando@info.unlp.edu.ar)
III-LIDI – School of Computer Science – UNLP - Argentina