Data Management for Big Data

Introduction to Part II Advanced database models, languages, and systems

Dario Della Monica

March 31, 2020

・ロト ・ 日 ・ モー・ モー・ うへぐ

Teacher's information

Dario Della Monica

Dipartimento di Scienze Matematiche, Informatiche e Fisiche (DMIF) University of Udine

- email: dario.dellamonica@uniud.it
- web: http://users.dimi.uniud.it/~dario.dellamonica/
 - ightarrow [Home] \gg Teaching \gg Data Management for Big Data
- office: room NN1, 2nd floor, nodo nord

Polo Scientifico "Rizzi" via delle Scienze, 206 33100 Udine

• office hours: Tuesday, 16:00-18:00 (or by appointment)

 Wednesday 16:00-18:00 when we have classes on Tuesday (next 3 weeks)

A ロ ト 4 同 ト 4 三 ト 4 三 ト 9 Q Q

- appointment is always preferable
- more info (cell, skype, ...) on my webpage

Material

- Database System Concepts Silberschatz, Korth, Sudarshan 7th Edition (or 6th Edition) McGraw-Hill, 2020 https://www.db-book.com/db7/index.html (extra contents: original slides, solutions, etc)
 - Principles of Distributed Database Systems
 Özsu and Valduriez
 3rd Edition
 Springer, 2011
 ISBN: 978-1-4939-4174-2
 http://extras.springer.com/ (original slides - use e-book ISBN: 978-1-4419-8834-8)
- other material (e.g., slides) is available on my webpage





▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ のへの

Overview of this part of the course

Data Management for Big Data

Part II: Advanced database models, languages, and systems

3 cfu (24 hours)

- 1. Query processing and optimization
 - Query processing
 - algorithms for evaluating operations (e.g., selection, join)
 - Cost-based optimization and heuristics
 - choice of "best" evaluation plan
- 2. Distributed database architectures
 - An introduction to distributed DBMS
 - Distributed database design (fragmentation and allocation)
 - Distributed query processing and optimization
 - Transaction management in distributed databases
 - distributed concurrency control (two-phase locking 2PL)

▲ロト ▲周ト ▲ヨト ▲ヨト ヨー のく⊙

- distributed DBMS reliability (two-phase commit 2PC)
- 3. Semistructured Data and XML
 - Definition of semistructured data in XML
 - Querying XML data (XPath and XQuery)
 - XML and relational DBMS
 - Native XML databases
- 4. Cloud computing and DBMS, Parallel DBMS