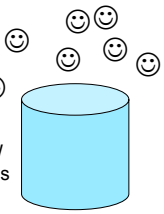
Data management in large-scale distributed systems

Claudia Lucia Roncancio

INP Grenoble
LIG – Sigma team
Claudia.Roncancio@imag.fr

Once upon a time…a DBMS!

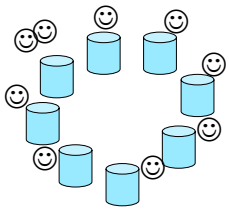
Models languages, DW, DM,…



data + R/W transactions

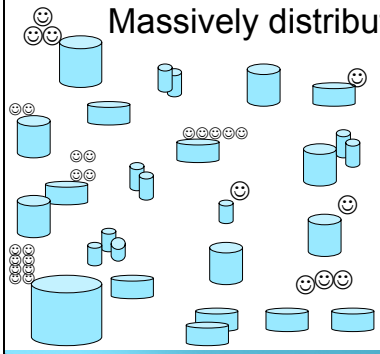
- Persistency
- Querying
- Transactions
- Concurrency control
- Fault tolerance
- Memory mgt. Caching
- Optimization

… distribution….



- Persistency
- Querying
- Transactions
- Concurrency control
- Fault tolerance
- Memory mgt. Caching
- Optimization

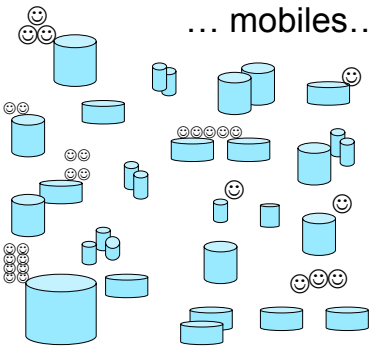
Massively distributed …



Heterogeneity ➶
Autonomy ➶
Dynamicity ➶

4

… mobiles…

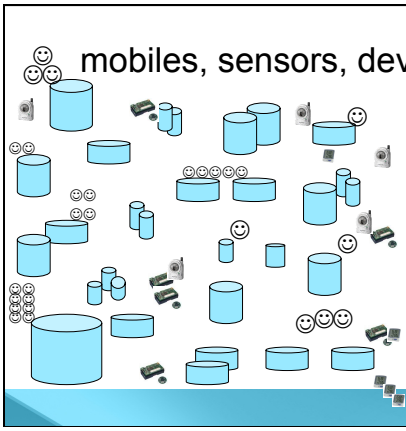


Heterogeneity ➶➶
Autonomy ➶➶
Dynamicity ➶➶

Context awareness

5

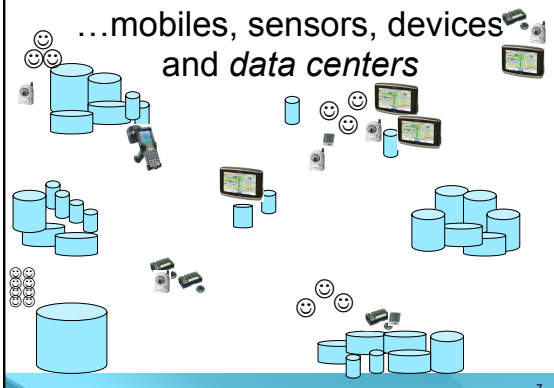
mobiles, sensors, devices …

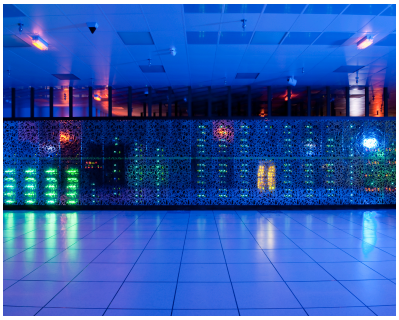


Heterogeneity ➶➶
Autonomy ➶➶
Dynamicity ➶➶

6

…mobiles, sensors, devices and *data centers*





Curie (supercalculateur Bull) en lumière, © CNRS Cyril Fresillon

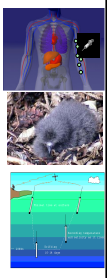
Distributed and ubiquitous systems

- Data everywhere!!
- Data, information and knowledge management
  - very small → big data
  - mobile, elastic, … !
- Extremely heterogeneous application domains and requirements

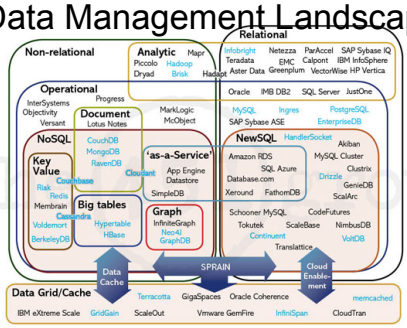
Application domains

Today the field covers all the largest sources of data ...

- Any company!
- Scientific and medical databases
- Integrating information
- Data mining
- Web search and data management
- Monitoring everywhere…

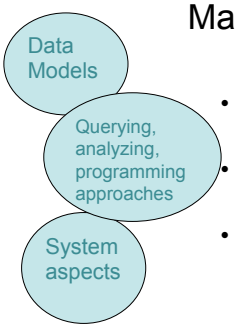


Data Management Landscape



From Matthew Aslett, http://blogs.the451group.com

Many issues



- Relational, Key-Value, Graphs, Document
- Map-Reduce, Spark frameworks
- Scalability, storage, concurrency control, consistency, transactions

…Dimensions of Big Data

Volume	Velocity	Variety	Veracity*
Data at Rest	Data in Motion	Data in Many Forms	Data in Doubt
Terabytes to exabytes of existing data to process	Streaming data, milliseconds to seconds to respond	Structured, unstructured, text, multimedia	Uncertainty due to data inconsistency & incompleteness, ambiguities, latency, deception, model approximations

* Truthfulness, accuracy or precision, correctness