HPI Hasso Plattner Institut
IT Systems Engineering | Universität Potsdam

Distributed Data Analytics
Introduction

Thorsten Papenbrock
G-3.1.09, Campus III
Hasso Plattner Institut

O'REILLY
Martin Kleppmann: *Designing Data-Intensive Applications*
O'Reilly Media, 2017
ISBN 978-1-4493-7332-0

Distributed Data Analytics

Information Systems Team

Thorsten **Papenbrock**

Diana **Stephan**

Prof. Felix **Naumann**

Dr. Ralf **Krestel**

Tim **Repke**

Sebastian **Kruse**

Julian **Risch**

Tobias **Bleifuß**

Hazar **Harmouch**

John **Koumarelas**

Michael **Loster**

Leon **Bornemann**

Zhe **Zuo**

Konstantina **Lazariduo**

Lan **Jiang**

Toni **Grütze**

Data Change
Data Fusion
Duplicate Detection
Data Profiling
Information Integration
Web Science
Data Scrubbing
Data as a Service
Data Cleansing
Information Quality
Text Mining
Web Data
RDF Data Mining
Dependency Detection
Linked Open Data
Service-Oriented Systems
ETL Management
Entity Recognition
Opinion Mining

project **DuDe**
project **Stratosphere**
project **DataChEx**
project **Metanome**

Distributed Data Analytics

Introduction

Introduction: Audience

Couses 2017/2018

Lectures

- Datenbanksysteme II (Bachelor)
- Data Engineering in der Praxis (Master/Bachelor)
- Information Retrieval and Web Search (Master)
- Information Integration (Master)
- Distributed Data Analytics (Master)

Seminars

- Beautiful Data (Bachelor)
- Probabilistic Graphical Models (Master)
- Advanced Data Profiling (Master)
- Building a Scalable Cryptocurrency with Big-Data Technology (Master)

This Lecture

Lecture

- For master students
- 3 credit points, 2 SWS
- Wednesdays 15:15 – 16:45

Exercises

- Interleaved with lectures during the semester

Slides

- On website

Website

- https://hpi.de/naumann/teaching/teaching/ws-1718/distributed-data-analytics-vl-master.html

Prerequisites

- To participate:
  A little background and interest in databases (e.g. DBS I lecture)
- For exam:
  Attending lectures, participation in exercises, and completion of exercise tasks

Exam

- Depending on number of students: oral or written exam
- Probably first week after lectures

Feedback

Official evaluation at end of semester

- This is a new lecture (and the first lecture for me ;-))

Question any time please!

- During lectures
- Visit me: Compus III, Room G-3.1.09
- Email: thorsten.papenbrock@hpi.de

Also: Give feedback about …

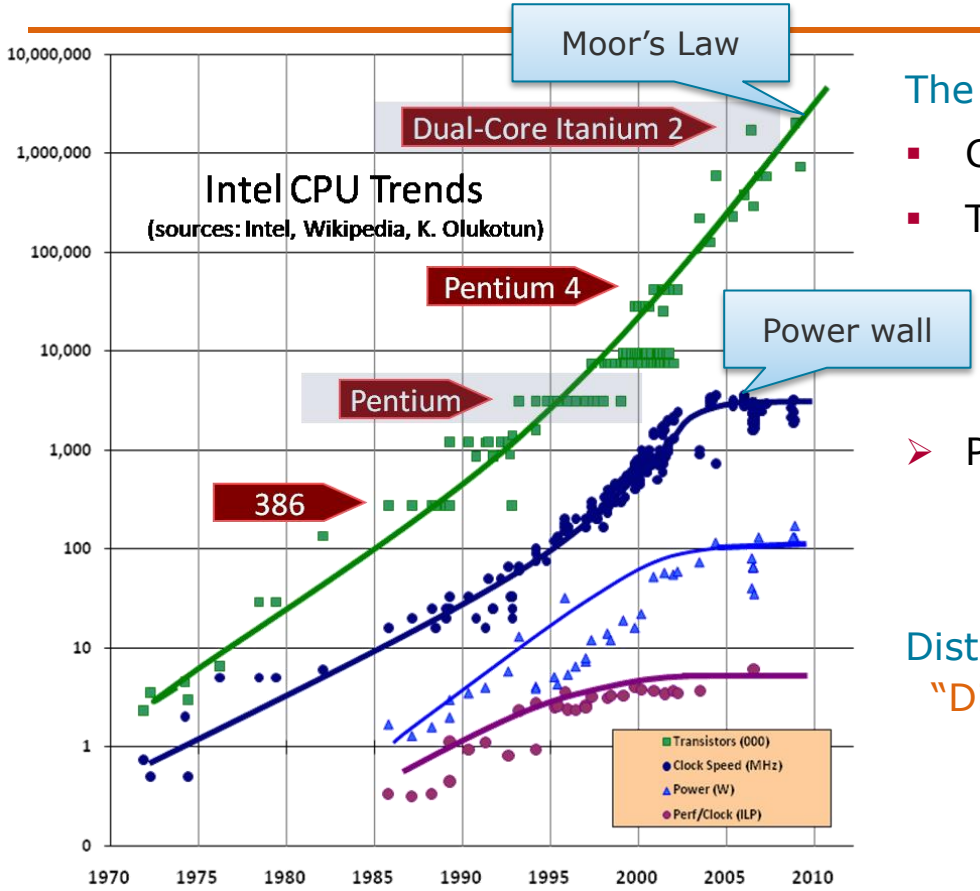
- improving lectures
- informational material
- organization

Distributed Data Analytics

Distributed Data Analytics

Motivation: "Distributed"

Paradigm Shift in Software-Writing



The free lunch is over!

- Clock speeds stall
- Transistor numbers still increase
  - Cores in CPUs/GPUs
    CPUs/GPUs in compute nodes,
    compute nodes in clusters
  - Paradigm Shift:
    - Earlier: optimize code for a single thread
    - Now: solve tasks in parallel

Distributed computing

"Distribution of work on (potentially) physically isolated compute nodes"

http://www.gotw.ca/publications/concurrency-ddj.htm

Motivation: "Distributed"

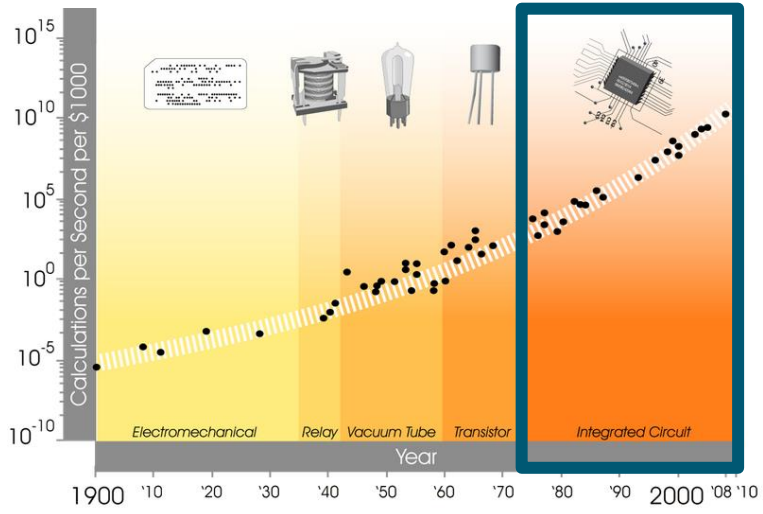
Surpassing Moor's Law

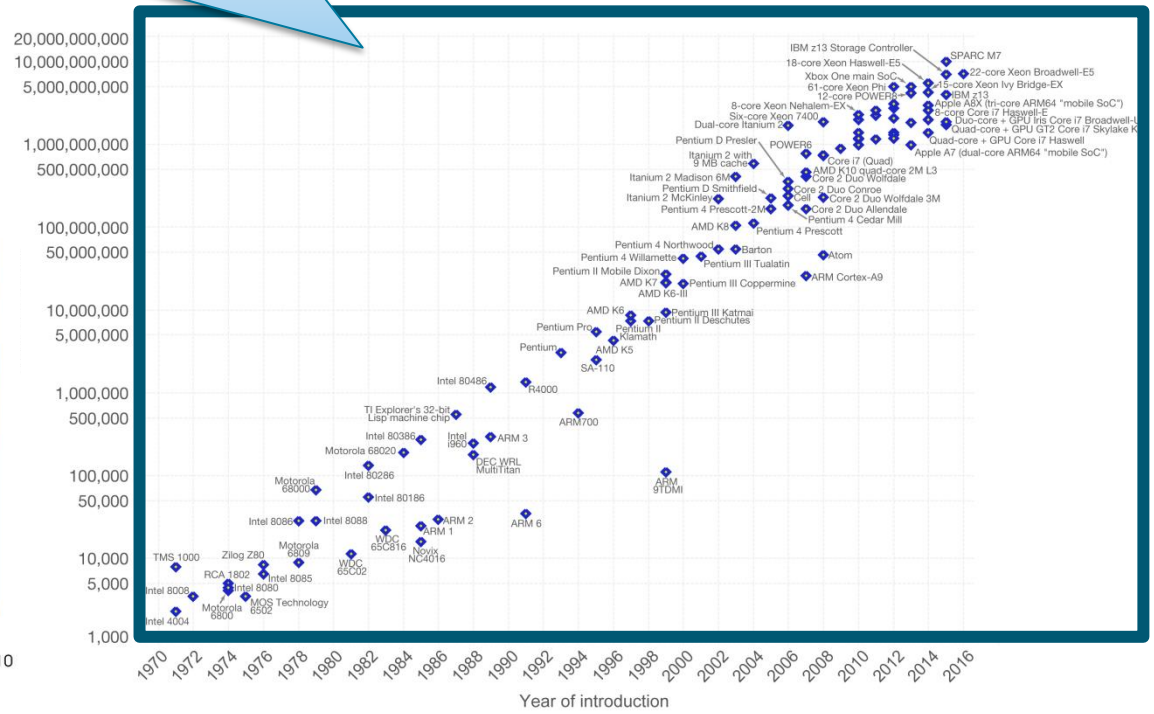
With **distributed machines**, we can already build systems with any number of transistors!

(don't even need to wait for a new processors)

Moore's Law (Observation)

"The number of transistors in integrated circuits doubles approximately every two years"





Motivation: "Distributed"

A Rule to Acknowledge

Amdahl's Law

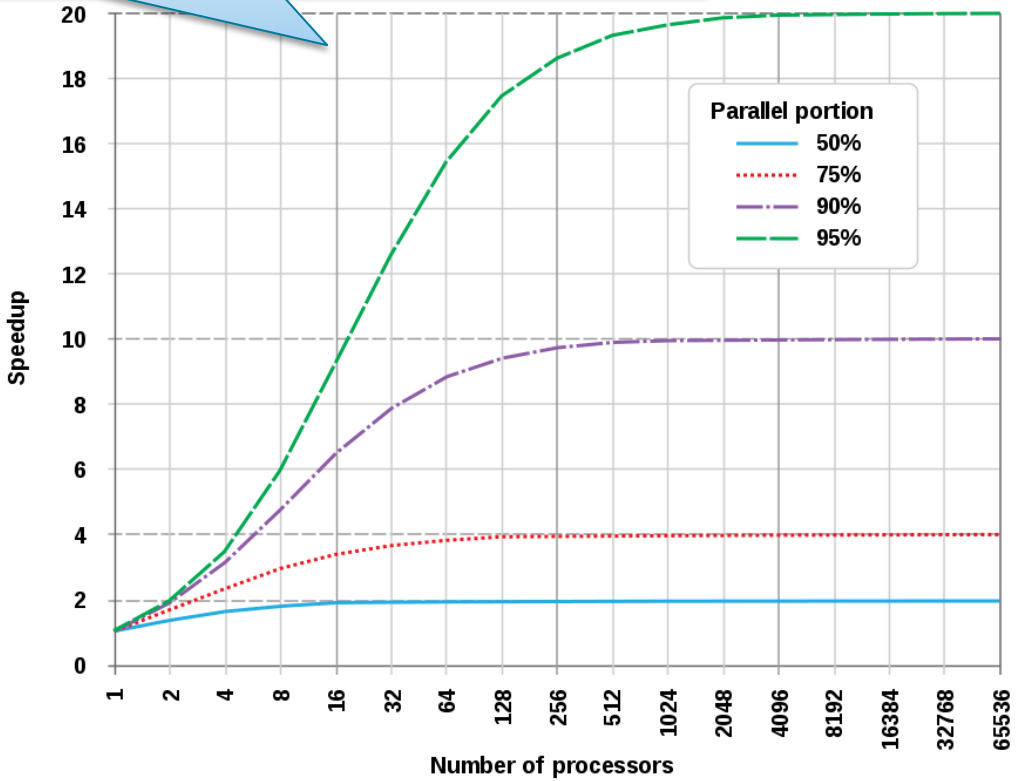
"The speedup of a program using multiple processors for parallel computing is limited by the sequential fraction of the program"

$$Speedup(s) = \frac{1}{(1-p) + \frac{p}{s}}$$

s: degree of parallelization (e.g. #cores)

p: percentage of the algorithm that profits from parallelization

Even **distributed parallelization** cannot work around this law!



New Technologies

Distributed Computing

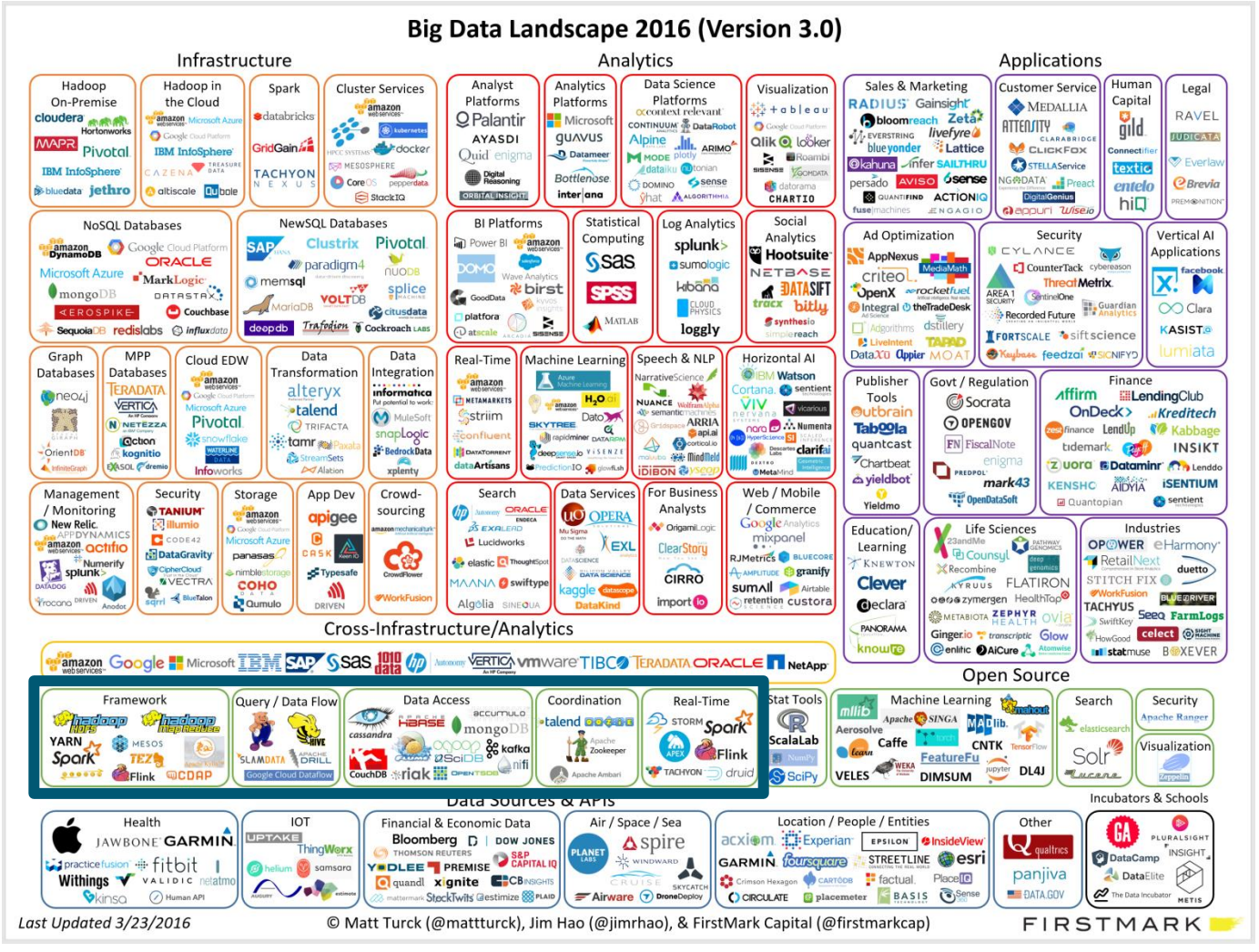
...

Distributed "Data Analytics"

Distributed Storage

"Distributed Data" Analytics

...



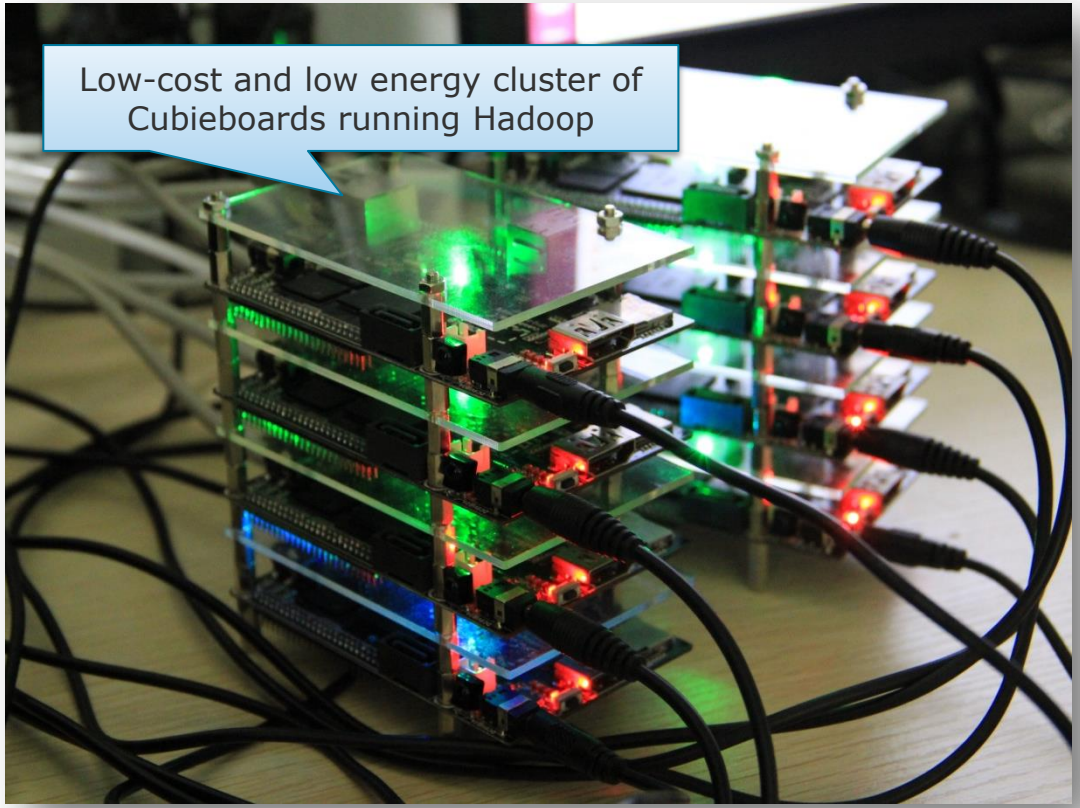
Hasso Plattner Institut

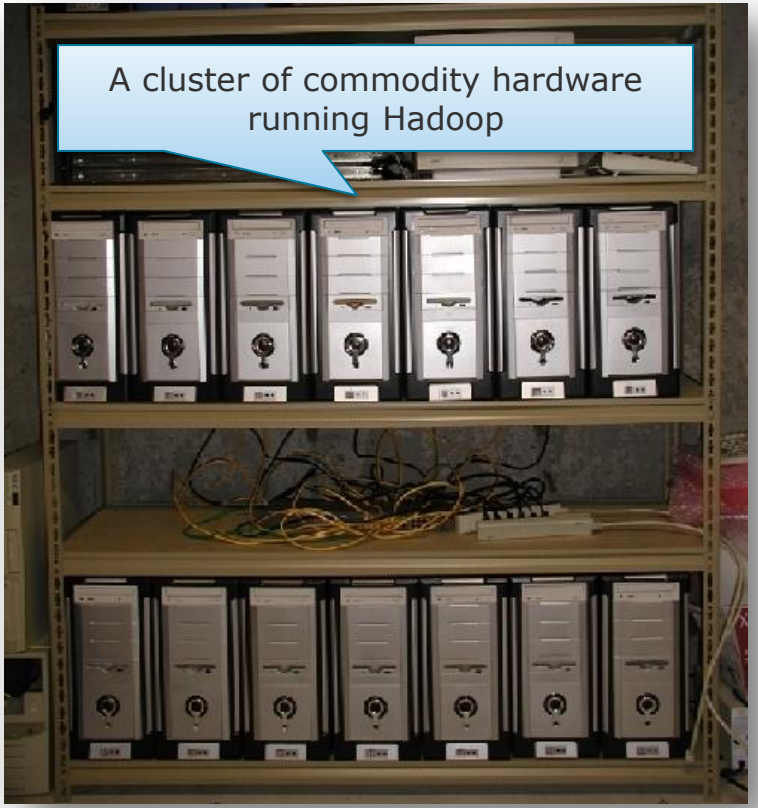
Distributed Data Analytics

Introduction

Motivation: “Distributed”

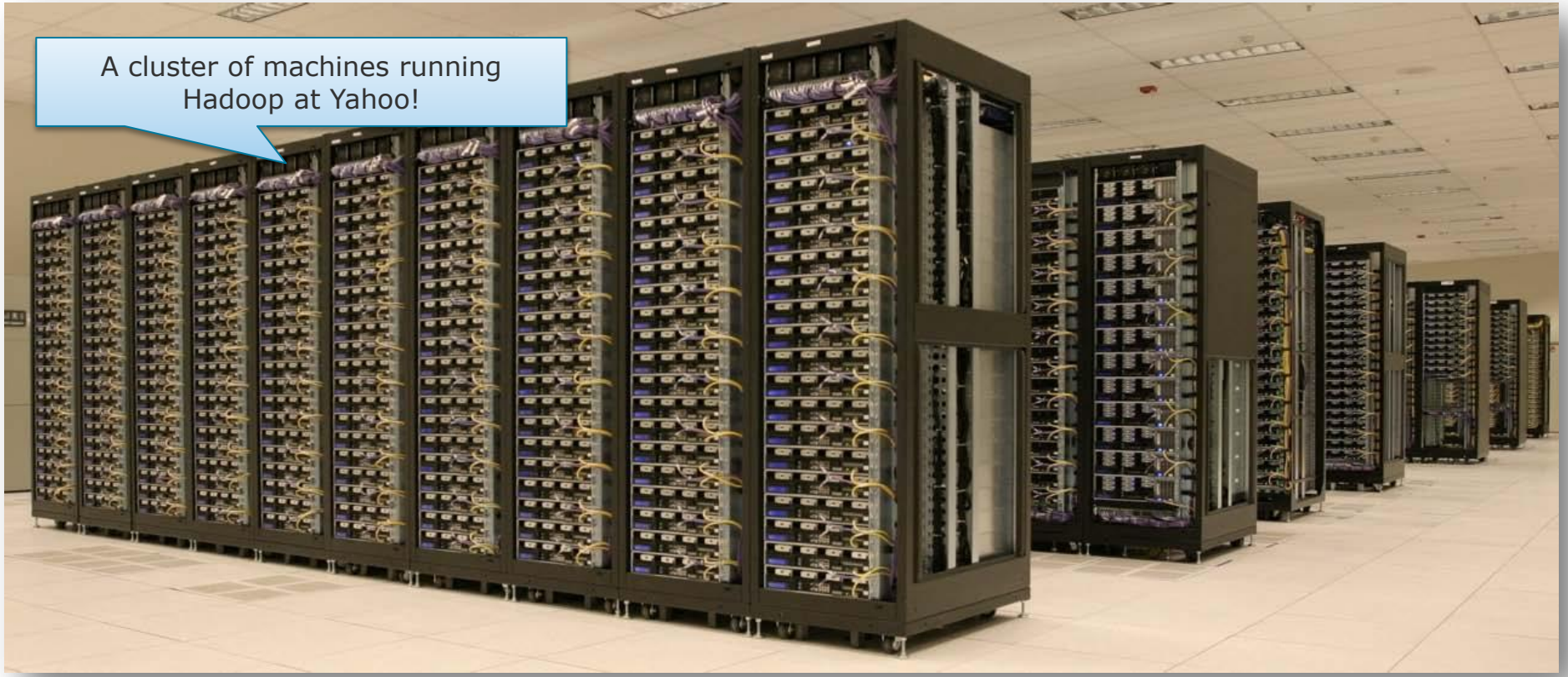
Small and Medium Scale





Motivation: "Distributed"

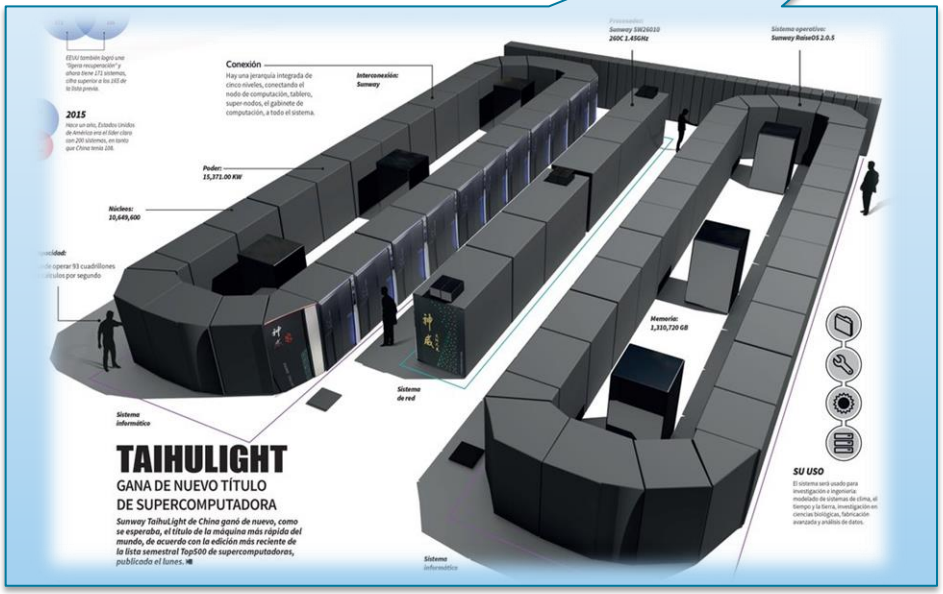
Large Scale



Motivation: "Distributed"

Super Large Scale

Top 10 Super Computers 2017



All **distributed systems**!

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
2	**Tianhe-2 (MilkyWay-2)** - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P , NUDT National Super Computer Center in Guangzhou China	3,120,000	33,862.7	54,902.4	17,808
3	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	361,760	19,590.0	25,326.3	2,272
4	**Titan** - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x , Cray Inc. DOE/SC/Oak Ridge National Laboratory United States	560,640	17,590.0	27,112.5	8,209
5	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LLNL United States	1,572,864	17,173.2	20,132.7	7,890
6	**Cori** - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/SC/LBNL/NERSC United States	622,336	14,014.7	27,880.7	3,939
7	**Oakforest-PACS** - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path , Fujitsu Joint Center for Advanced High Performance Computing Japan	556,104	13,554.6	24,913.5	2,719
8	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect , Fujitsu RIKEN Advanced Institute for Computational Science (AICS) Japan	705,024	10,510.0	11,280.4	12,660
9	**Mira** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom , IBM DOE/SC/Argonne National Laboratory United States	786,432	8,586.6	10,066.3	3,945
10	**Trinity** - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect , Cray Inc. DOE/NNSA/LANL/SNL United States	301,056	8,100.9	11,078.9	4,233

https://www.top500.org/lists/2017/06/

Motivation: "Distributed"

Super Large Scale

Top 10 Super Computers 2017

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
2	**Tianhe-2 (MilkyWay-2)** - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P , NUDT National Super Computer Center in Guangzhou China	3,120,000	33,862.7	54,902.4	17,808
3	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	361,760	19,590.0	25,326.3	2,272
4	**Titan** - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x , Cray Inc. DOE/SC/Oak Ridge National Laboratory United States	560,640	17,590.0	27,112.5	8,209
5	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LLNL United States	1,572,864	17,173.2	20,132.7	7,890
6	**Cori** - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/SC/LBNL/NERSC United States	622,336	14,014.7	27,880.7	3,939
7	**Oakforest-PACS** - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path , Fujitsu Joint Center for Advanced High Performance Computing Japan	556,104	13,554.6	24,913.5	2,719
8	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect , Fujitsu RIKEN Advanced Institute for Computational Science (AICS) Japan	705,024	10,510.0	11,280.4	12,660
9	**Mira** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom , IBM DOE/SC/Argonne National Laboratory United States	786,432	8,586.6	10,066.3	3,945
10	**Trinity** - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect , Cray Inc. DOE/NNSA/LANL/SNL United States	301,056	8,100.9	11,078.9	4,233

https://www.top500.org/lists/2017/06/

All **distributed systems**!

Motivation: "Distributed"

Super Large Scale

Top 10 Super Computers 2017

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
2	**Tianhe-2 (MilkyWay-2)** - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P , NUDT National Super Computer Center in Guangzhou China	3,120,000	33,862.7	54,902.4	17,808
3	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	361,760	19,590.0	25,326.3	2,272
4	**Titan** - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x , Cray Inc. DOE/SC/Oak Ridge National Laboratory United States	560,640	17,590.0	27,112.5	8,209
5	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LLNL United States	1,572,864	17,173.2	20,132.7	7,890
6	**Cori** - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/SC/LBNL/NERSC United States	622,336	14,014.7	27,880.7	3,939
7	**Oakforest-PACS** - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path , Fujitsu Joint Center for Advanced High Performance Computing Japan	556,104	13,554.6	24,913.5	2,719
8	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect , Fujitsu RIKEN Advanced Institute for Computational Science (AICS) Japan	705,024	10,510.0	11,280.4	12,660
9	**Mira** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom , IBM DOE/SC/Argonne National Laboratory United States	786,432	8,586.6	10,066.3	3,945
10	**Trinity** - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect , Cray Inc. DOE/NNSA/LANL/SNL United States	301,056	8,100.9	11,078.9	4,233

https://www.top500.org/lists/2017/06/

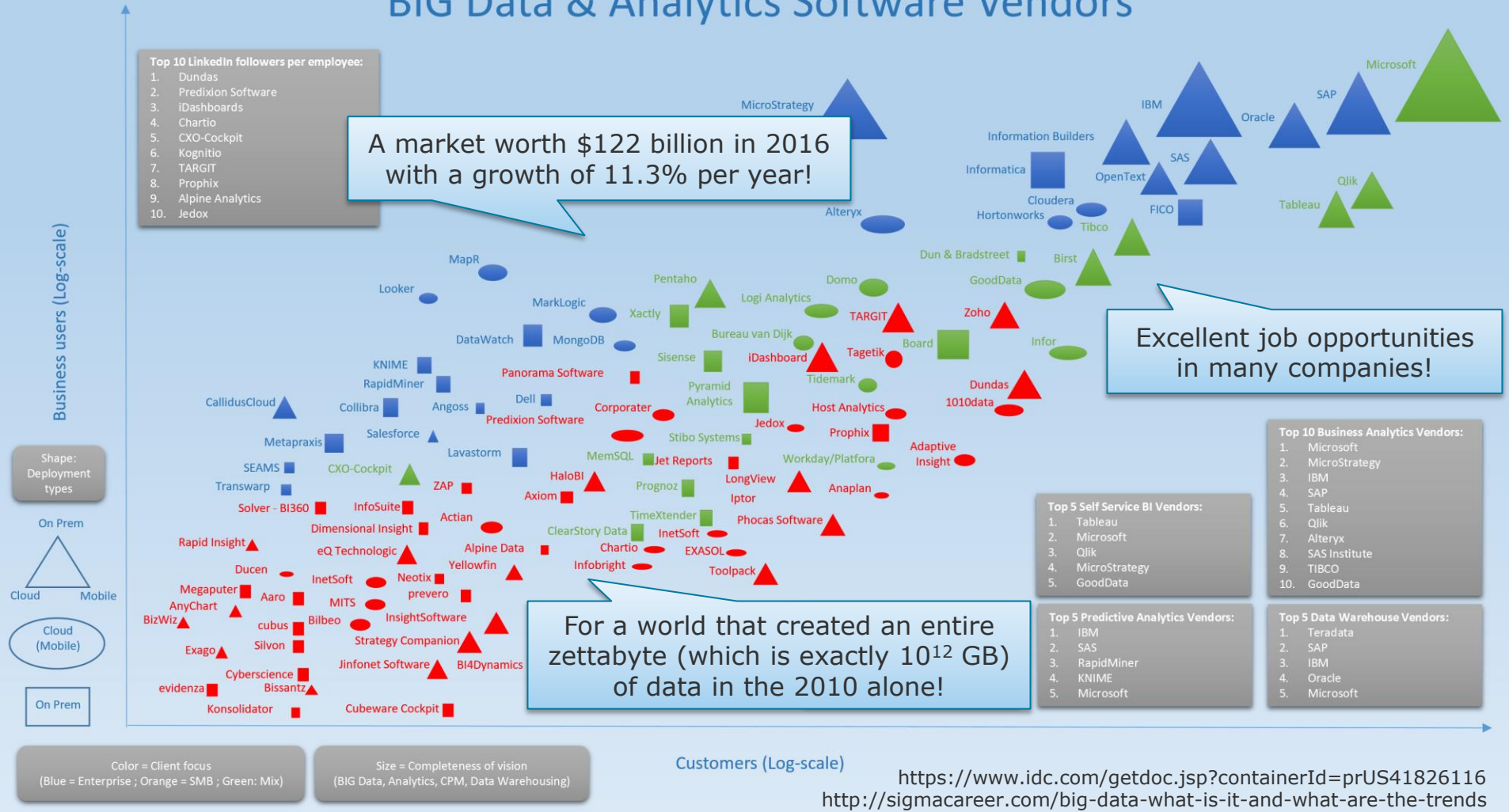
All **distributed systems**!

Distributed Data Analytics

"The ability of extracting different kinds of information from data!"

- Structural information
- Explicit information
- Implicit/derived information

BIG Data & Analytics Software Vendors



https://www.idc.com/getdoc.jsp?containerId=prUS41826116
http://sigmacareer.com/big-data-what-is-it-and-what-are-the-trends

VLDB 2017 Program

International conference "Very Large Data Bases"

29 Aug (Tue)								
8:30-10:00		Welcome Messages + Wolfgang Lehner keynote						
10:00-10:30	break							
10:30-12:00		Panel		DB Engines 1	Data Cleaning	Spatial Data Management 1	Graphs and Networks 1	Tutorial: Caching at the Web Scale
12:00-13:30	lunch							
13:30-15:00		Query Processing and Optimization		Stream Processing 1	Data Formats	Privacy and Security	Event Processing	Tutorial: Summarizing Static and Dynamic Big Graphs
15:00-15:30	break							
:00		Transactions		Spatial Data management 2	Graphs and Networks 2	Information Integration	Applications	Tutorial: Blockchains and Databases
:00			poster reception					
(Wed)								
8:30-10:00		Endowment Update + VLDB Awards + Michael Franklin keynote						
10:00-10:30	break							
10:30-12:00		High Performance Query Processing		Recommendations and Skylines	Approximation Structures	Stream Processing 2		Tutorial: New Trends on Exploratory Methods for Data Analytics
12:00-13:30	lunch							
13:30-15:00		Scalable Storage		Crowdsourcing	Stream Processing 3	Graphs and Networks 3	Social Network Analysis	Tutorial: Geometric Approaches for Top-k Queries
15:00-15:30	break							
15:30-17:00		Concurrency Control		Query Optimization	Specialized Data Management	Graphs and Networks 4	Data Mining and Analytics	Tutorial: Spatial Crowdsourcing: Challenges, Techniques, and Applications
17:00-18:00			poster reception					
19:00 - ...	**"VLDB Octoberfest" banquet at the Hofbräuhaus**							
31 Aug (Thu)								
8:30-10:30		Endowment Awards + Demo Award + Jens Dittrich plenary						
10:30-10:55	break							
10:55-12:00		Transactions and Persistence		Data Access	Data Statistics	Potpourri		
12:00-13:30	lunch							
13:30-15:00		Data Partitioning		Graphs and Networks 5	Visualization	Distributed Systems and Cloud 1	Estimation and Approximation	Tutorial: The Era of Big Spatial Data
15:00-15:30	break							
15:30-17:00		DB Engines 2		Query Processing	Text and Semi-Structured	Spatial Data Management 3	Distributed Systems and Cloud 2	Tutorial: Complex Event Recognition in the Big Data Era
17:00-18:00			poster reception					

VLDB 2017 Program

29 Aug (Tue)								
8:30-10:00		Welcome Messages + Wolfgang Lehner keynote						
10:00-10:30	break							
10:30-12:00		Panel		DB Engines 1	Data Cleaning	Spatial Data Management 1	Graphs and Networks 1	Tutorial: Caching at the Web Scale
12:00-13:30	lunch							
13:30-15:00		Query Processing and Optimization		Stream Processing 1	Data Formats	Privacy and Security	Event Processing	Tutorial: Summarizing Static and Dynamic Big Graphs
15:00-15:30	break							
15:30-17:00		Transactions		Spatial Data management 2	Graphs and Networks 2	Information Integration	Applications	Tutorial: Blockchains and Databases
17:00-18:00			poster reception					
30 Aug (Wed)								
8:30-10:00		Endowment Update + VLDB Awards + Michael Franklin keynote						
10:00-10:30	break							
10:30-12:00		High Performance Query Processing		Recommendations and Skylines	Approximation Structures	Stream Processing 2		Tutorial: New Trends on Exploratory Methods for Data Analytics
12:00-13:30	lunch							
13:30-15:00		Scalable Storage		Crowdsourcing	Stream Processing 3	Graphs and Networks 3	Social Network Analysis	Tutorial: Geometric Approaches for Top-k Queries
15:00-15:30	break							
15:30-17:00		Concurrency Control		Query Optimization	Specialized Data Management	Graphs and Networks 4	Data Mining and Analytics	Tutorial: Spatial Crowdsourcing: Challenges, Techniques, and Applications
17:00-18:00			poster reception					
19:00-		**"VLDB Octoberfest" banquet at the Hofbräuhaus**						
31 Aug (Thu)								
8:30-10:30		Endowment Awards + Demo Award + Jens Dittrich plenary						
10:30-10:55	break							
10:55-12:00		Transactions and Persistence		Data Access	Data Statistics	Potpourri		
12:00-13:30	lunch							
13:30-15:00		Data Partitioning		Graphs and Networks 5	Visualization	Distributed Systems and Cloud 1	Estimation and Approximation	Tutorial: The Era of Big Spatial Data
15:00-15:30	break							
15:30-17:00		DB Engines 2		Query Processing	Text and Semi-Structured	Spatial Data Management 3	Distributed Systems and Cloud 2	Tutorial: Complex Event Recognition in the Big Data Era
17:00-18:00			poster reception					

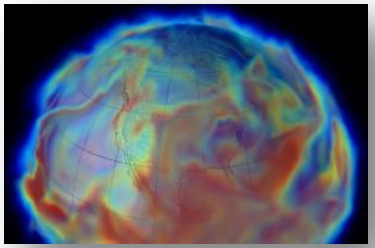
With strong relation to **data analytics** and/or **distributed processing**

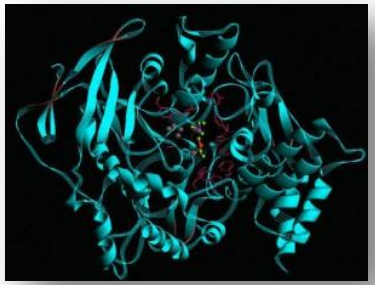
Motivation: "Data Analytics"

Back to Super Computers

For what tasks are these computer used?

- Weather forecasting
- Market analysis
- Crash simulation
- Disaster simulation
- Brute force decryption
- Molecular dynamics modeling
- …





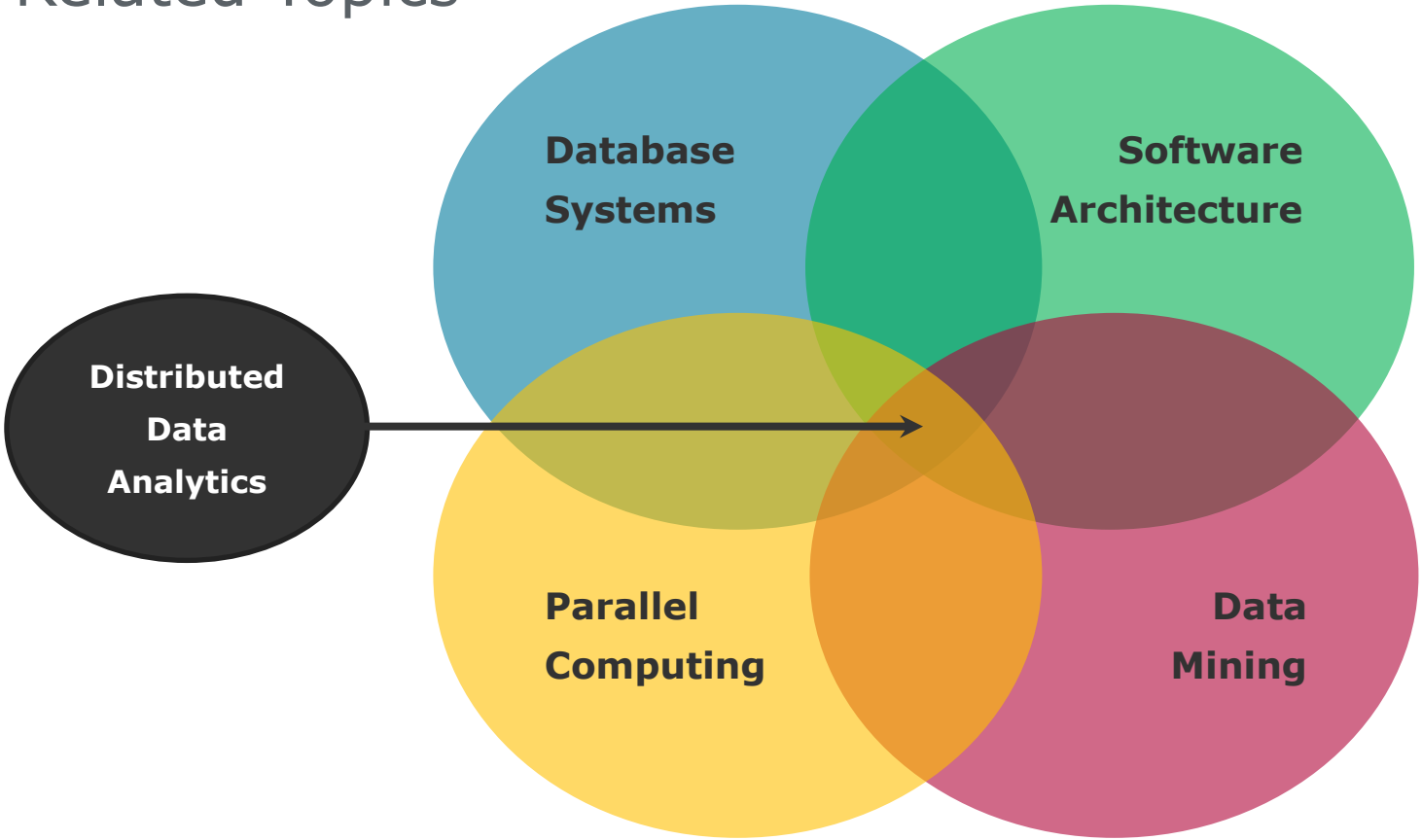
Data Analytics tasks!

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
2	**Tianhe-2 (MilkyWay-2)** - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P , NUDT National Super Computer Center in Guangzhou China	3,120,000	33,862.7	54,902.4	17,808
3	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	361,760	19,590.0	25,326.3	2,272
4	**Titan** - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x , Cray Inc. DOE/SC/Oak Ridge National Laboratory United States	560,640	17,590.0	27,112.5	8,209
5	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LLNL United States	1,572,864	17,173.2	20,132.7	7,890
6	**Cori** - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/SC/LBNL/NERSC United States	622,336	14,014.7	27,880.7	3,939
7	**Oakforest-PACS** - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path , Fujitsu Joint Center for Advanced High Performance Computing Japan	556,104	13,554.6	24,913.5	2,719
8	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect , Fujitsu RIKEN Advanced Institute for Computational Science (AICS) Japan	705,024	10,510.0	11,280.4	12,660
9	**Mira** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom , IBM DOE/SC/Argonne National Laboratory United States	786,432	8,586.6	10,066.3	3,945
10	**Trinity** - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect , Cray Inc. DOE/NNSA/LANL/SNL United States	301,056	8,100.9	11,078.9	4,233

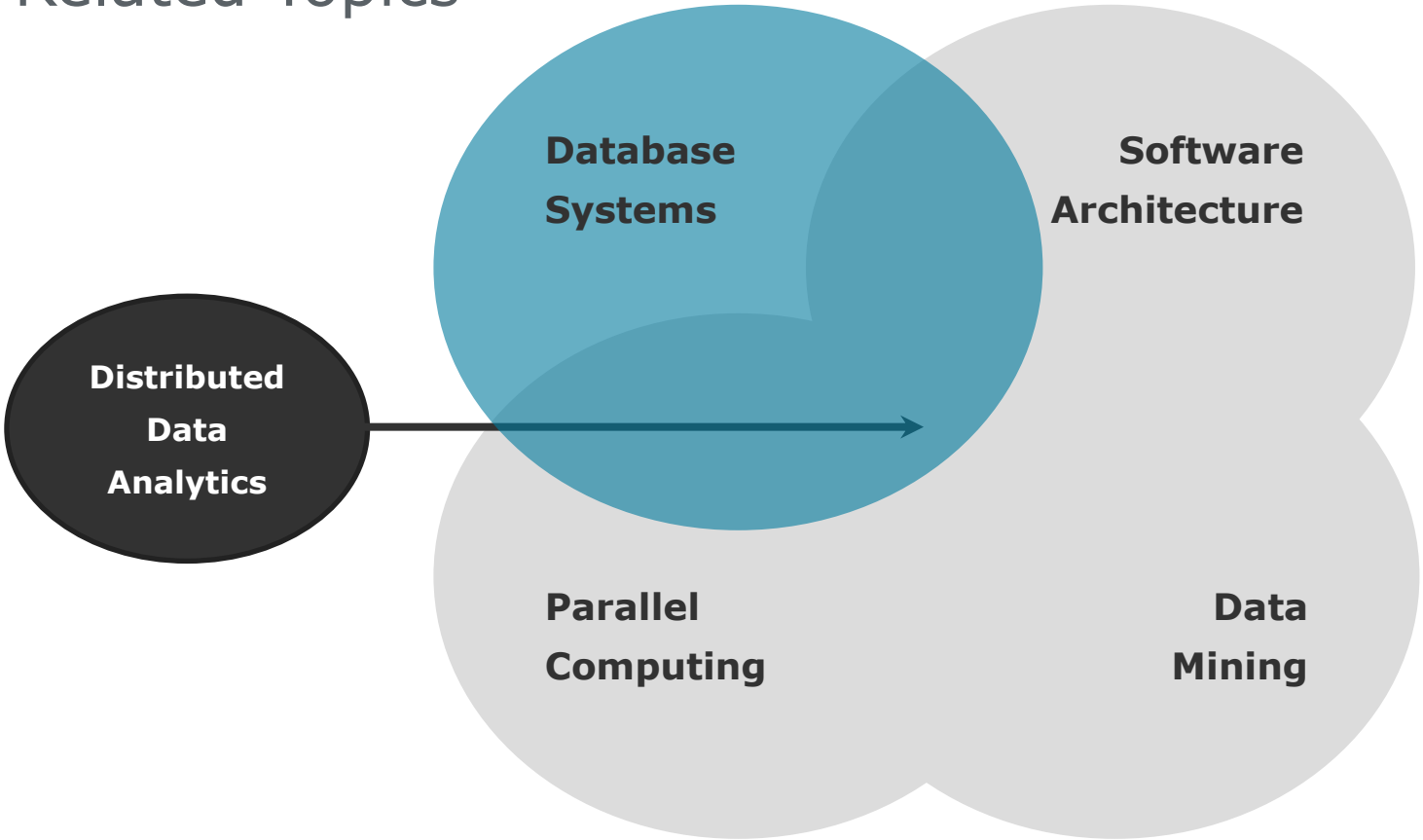
Driving Forces

- Data volumes increase:
  business data, sensor data, social media data, …
- Data analytics gains importance:
  downtime-less, real-time, predictive
- Parallelization paradigm shifts:
  multi-core and network speeds increase while CPU clock speeds stall
- Computation resources become more available:
  IaaS, PaaS, SaaS
- Free and open source software gains popularity:
  setting standards, utilizing external development resources, improving software quality, avoiding vendor locks …

Related Topics



Related Topics



Database Systems

Touch points

- Data models, query languages, and consistency guarantees
- Distributed storage and retrieval of data
- Index structures

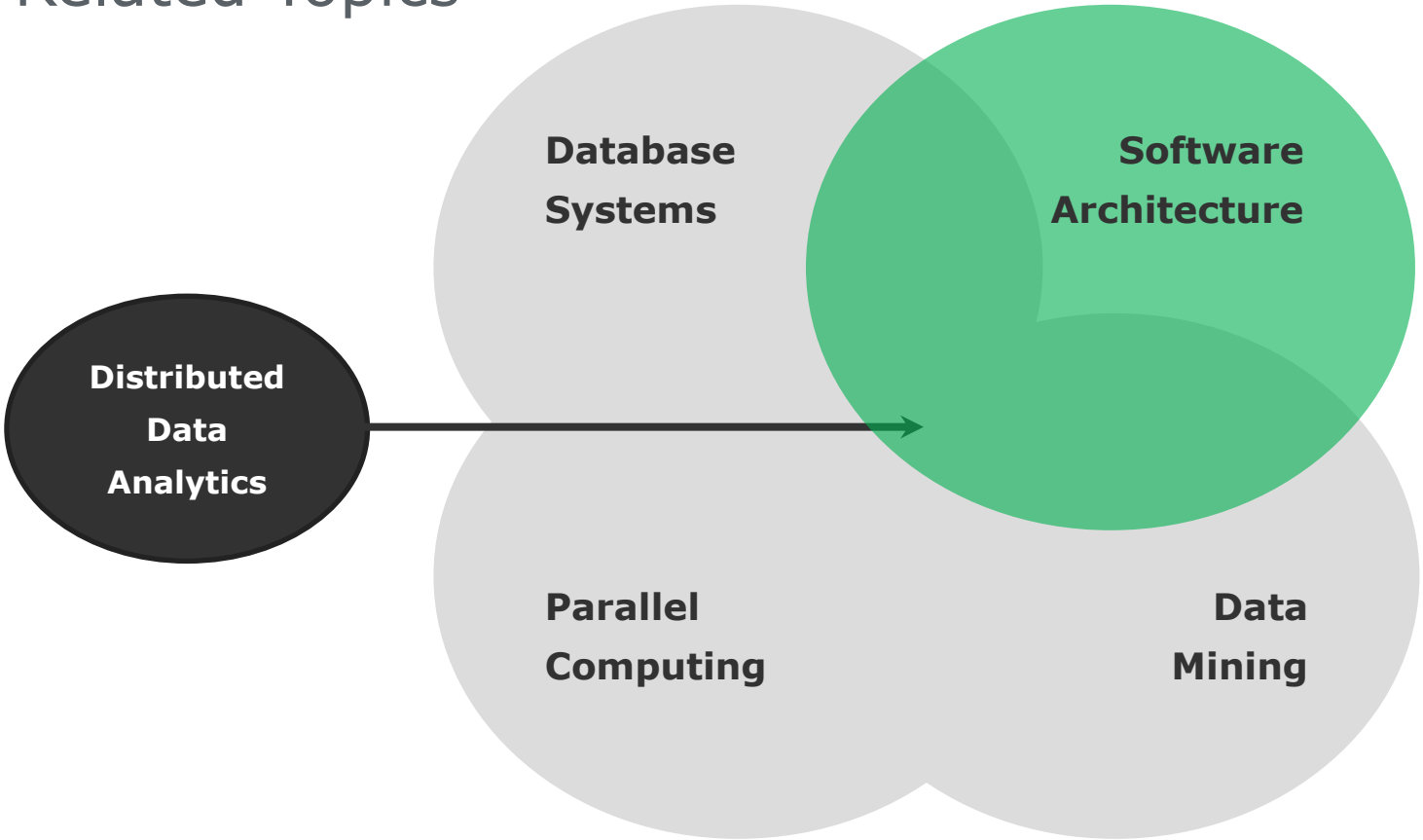
Not in this lecture

- Physical data storage
- In-depth presentations language constructs and transactions
- Core database technology, e.g., query optimizer

More focused lectures

- Database Systems I + II (Prof. Naumann)
- Trends and Concepts in Software Industry (Prof. Plattner)

Related Topics



Software Architectures

Touch points

- Requirements, design, and architecture of **distributed analytical** systems
- Pros and cons of different technologies for **distributed analytical** systems

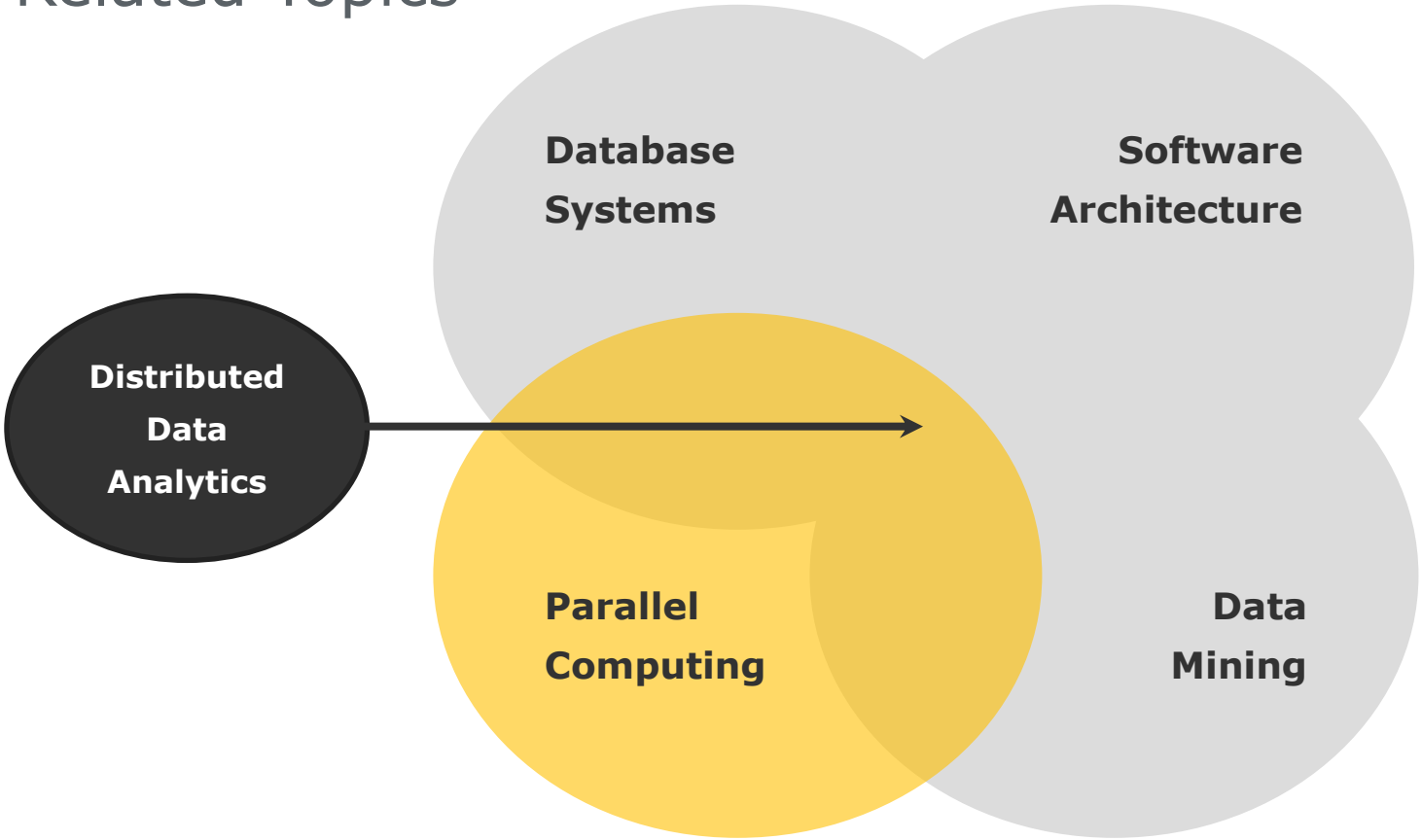
Not in this lecture

- Non-distributed or non-analytical systems or services
- Agile software development techniques
- Software patterns

More focused lectures

- Software Architecture (Dr. Ulflacker)
- Software Technique (Dr. Ulflacker)

Related Topics



Parallel Computing

Touch points

- Distributed data storage concepts
- Distributed programming models, e.g., actor programming and MapReduce

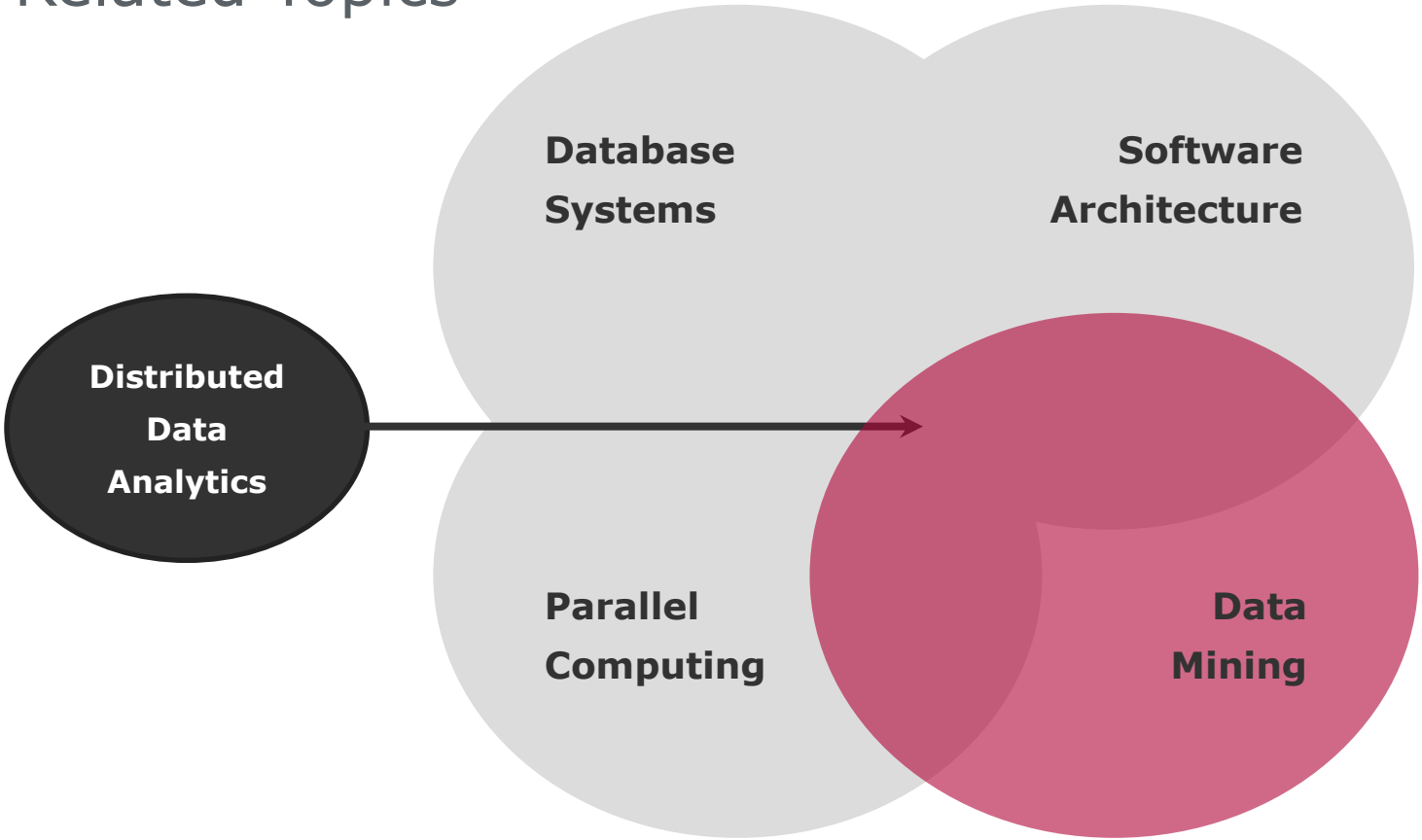
Not in this lecture

- Parallel, non-distributed programming languages, e.g., CUDA or OpenMP
- Core parallel computing concepts, e.g., scheduling or shared memory
- Processor architectures, cache hierarchies, GPU programming, …

More focused lectures

- Parallel Programming (Dr. Tröger)
- Programmierung paralleler und verteilter Systeme (Dr. Feinbube)

Related Topics



Data Mining

Touch points

- Data analytics: aggregation queries and basic data mining algorithms

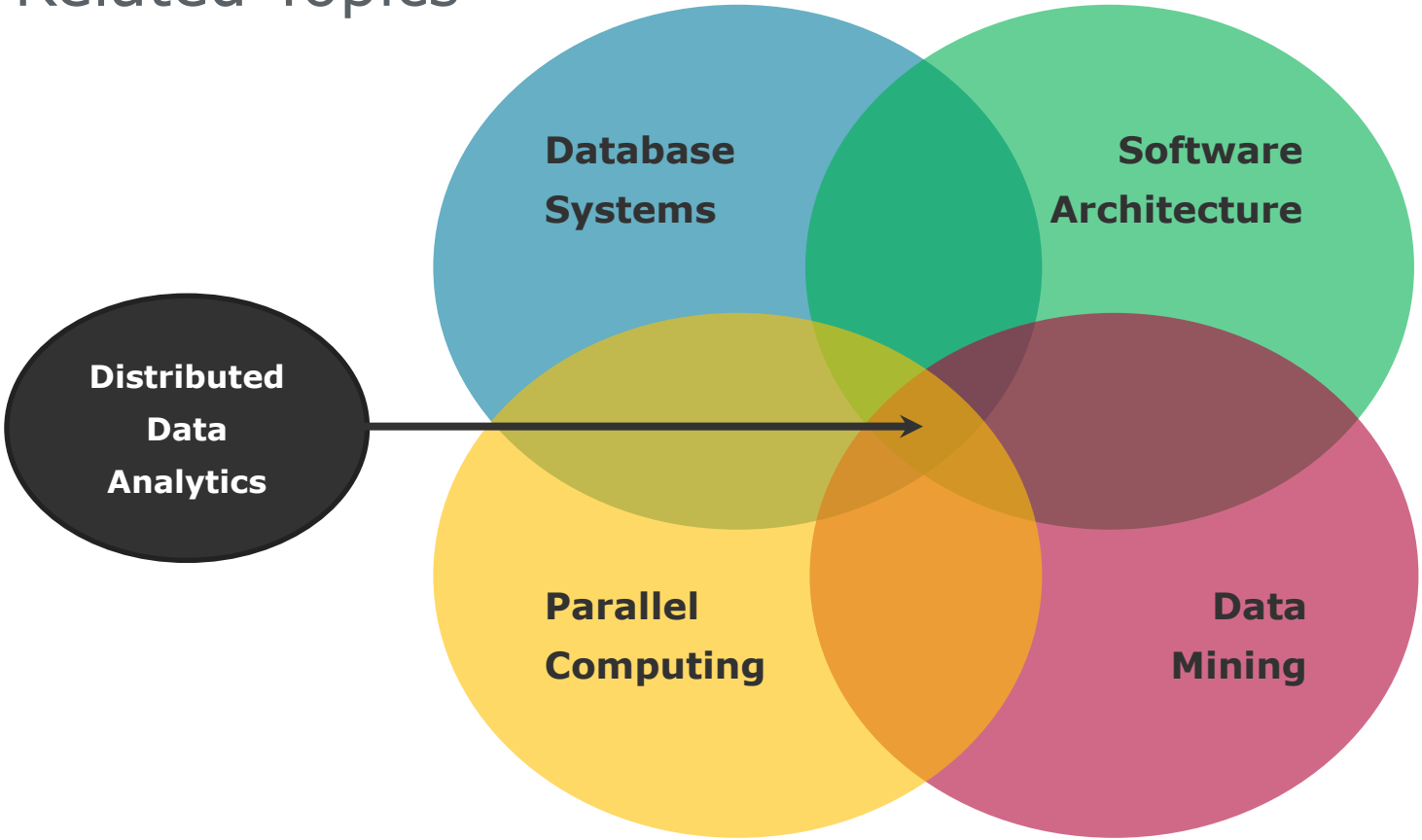
Not in this lecture

- Detailed introduction to machine learning, e.g., neuronal networks, (un)supervised learning, or Bayesian classification
- Statistics, linear algebra, and most sophisticated mining algorithms

More focused lectures

- Big Data Analytics (Prof. Müller)
- Graph Mining (Prof. Müller)
- Data Profiling (Prof. Naumann)
- Data Mining and Probabilistic Reasoning (Dr. Krestel)

Related Topics



Lecture Goals

Sorting the buzzwords

- NoSQL, Big Data, OLAP, Web-scale, ACID, Sharding, MapReduce, …

Understanding the systems

- How and why do distributed systems work?
- What are their individual core technologies?
- Which advantages and disadvantages do the technologies have?
- How can different technologies be combined?

Exercising in data analytics

- Distributed processing of data-parallel tasks
- Data mining on streams

Lecture Outline

Schedule

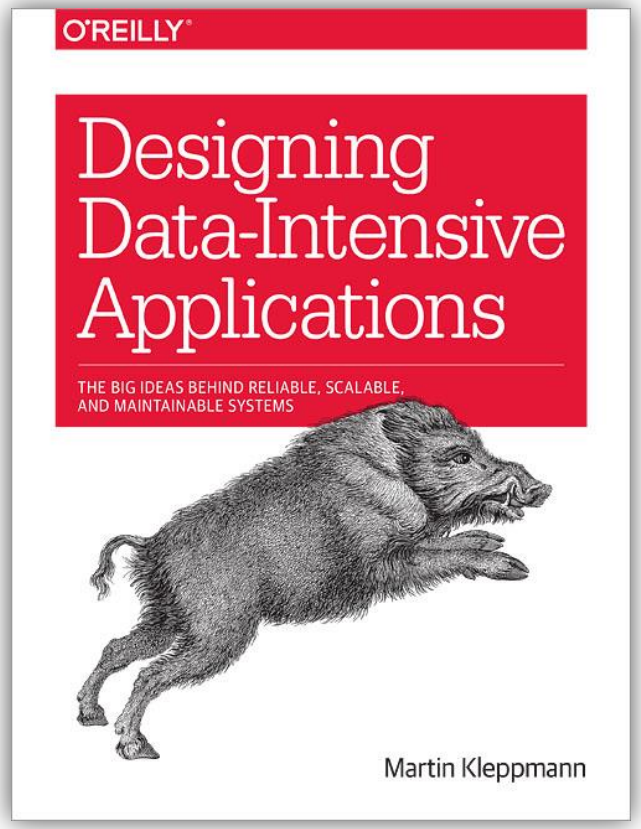
Date	Topic
18.10.2017	Introduction
25.10.2017	Foundations
01.11.2017	Data Models and Query Languages
08.11.2017	Storage and Retrieval
15.11.2017	Encoding and Evolution
22.11.2017	Hands-On: Akka
29.11.2017	Replication
06.12.2017	Partitioning
13.12.2017	Transactions
20.12.2017	Batch Processing
03.01.2018	Hands-On: Spark
10.01.2018	Stream Processing
17.01.2018	Distributed Systems
24.01.2018	Consistency and Consensus
31.01.2018	
07.02.2018	

Exercise 1

Exercise 2

Exam

Literature: Course Book



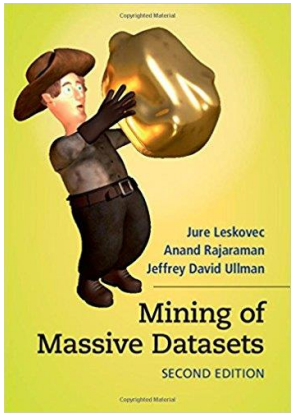
Designing Data-Intensive Applications

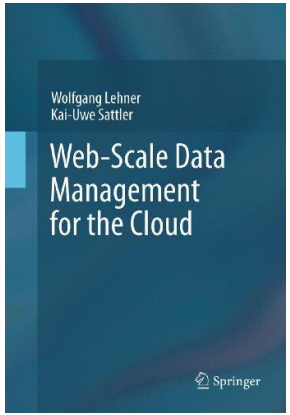
- Author: Martin Klappmann
- Date: March 2017
- Publisher: O'Reilly Media, Inc
- ISBN: 978-1-449-37332-0
- References: https://github.com/ept/ddia-references

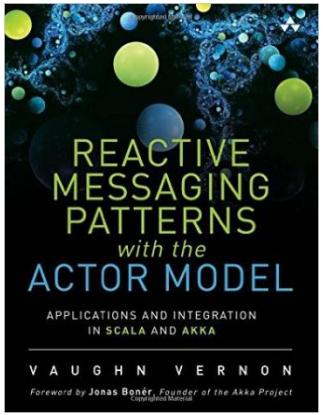
Scope for this lecture

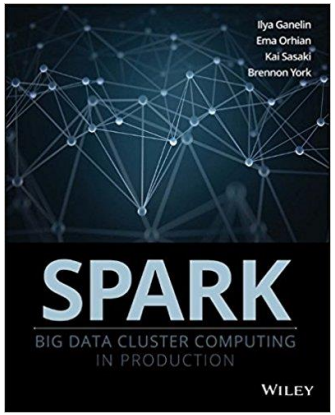
- Distributed and parallel systems for data analytics and data mining
- Big data storage
- Batch and stream processing

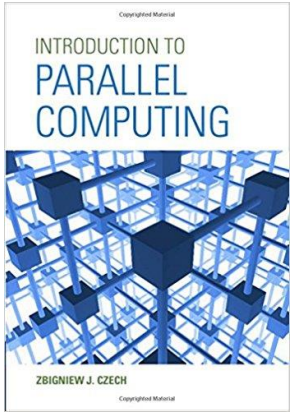
Literature: Further Reading

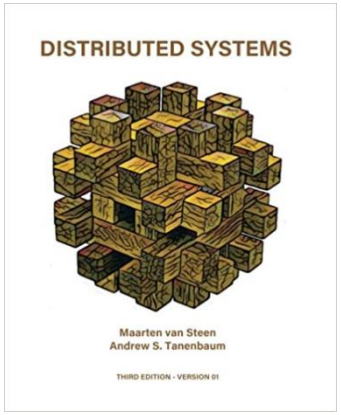


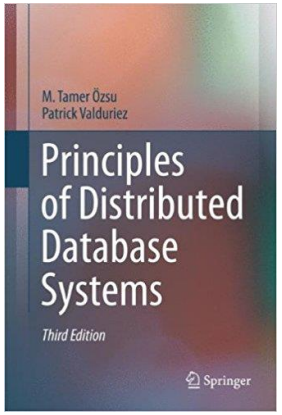




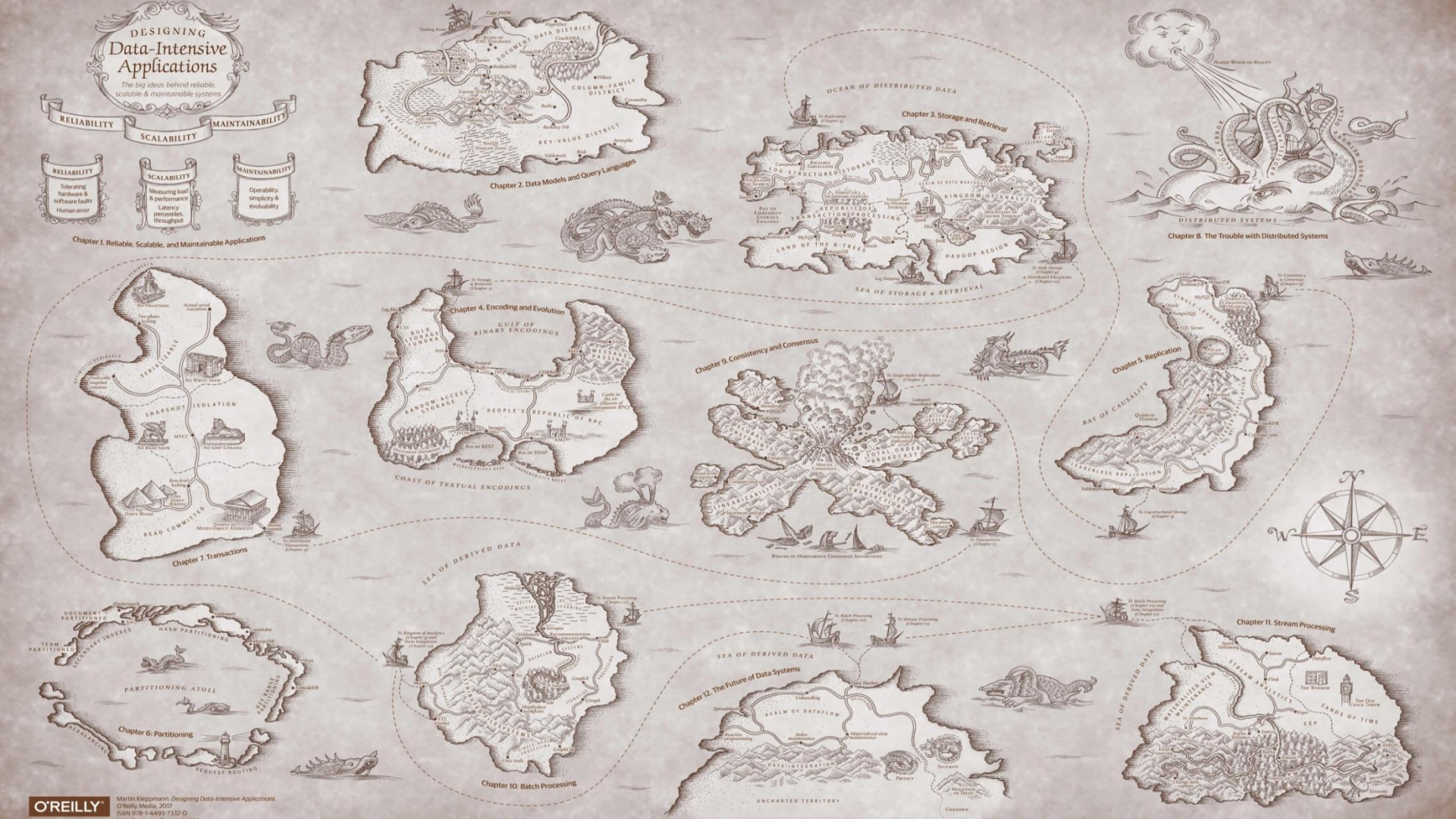








Web-links are given on the slides during the lecture.



O'REILLY
Martin Kleppmann: *Designing Data-Intensive Applications*
O'Reilly Media, 2017
ISBN 978-1-4493-7332-0