Entwicklung einer Lern-Feedback-Plattform für Vorlesungen

Lecture Monitoring - LeMon

Studienarbeit

Semester 5 + 6

Studiengang *Informatik*

Vertiefungsrichtung *Angewandte Informatik*

an der
Dualen Hochschule Baden-Württemberg / Stuttgart

von

Ephraim Petry & Nick Herrmannsdörfer

06. Juni 2014

Bearbeitungszeitraum	zwei Semester
Matrikelnummer Ephraim Petry, Kurs	2767400, TINF11D
Matrikelnummer Nick Herrmannsdörfer, Kurs	1655361, TINF11D
Betreuer der Praxisphase	Frau Dr. Barbara Dörsam

Zusammenfassung

Im Rahmen dieser Studienarbeit sollte ein System erstellt werden, mit dem der Lernfortschritt und das Verständnis der Studenten an der für die in der Vorlesung vermittelten Inhalte abgeprüft werden sollte. Dieses System sollte heißen.

Das System sollte dabei den kompletten Prozess von der Erstellung einzelner Übungsblätter durch den Dozenten, das Management der Fragen für die Übungsbläter, sowie das Freischalten der Übungsblätter zur entsprechenden Zeit beinhalten. Desweiteren sollte der Dozent den Studenten auf einfache Art und Weise das elektronische Übungsblatt zugänglich machen. Hierzu sollte bevorzugt ein QR-Code eingesetzt werden, den die Studenten mit ihrem Smartphone und einer entsprechenden App sofort einlesen und zum Übungsblatt gelangen konnten.
Am Ende sollte der Dozent auf einfache Weise die Möglichkeit haben, die Ergebnisse wenige Sekunden nach Ablauf der Zeit anschaulich angezeigt zu bekommen und diese mit den Studenten durchsprechen.

Zur Umsetzung dieses Projektes sollte eng mit einer Gruppe mit einem ähnlichen Projekt zusammengearbeitet werden, die jedoch eine etwas andere Art von Übungsblättern für den Einsatz von Aufgaben innerhalb oder auch außerhalb der Vorlesung erstellen sollte. Soweit möglich sollten dabei Synergien genutzt und eine gemeinsame Oberfläche für beide Projekte erstellt werden.

Contents

1 Einleitung **4**
1.1 Überschrift abc 4

2 Use Cases **5**

3 Fazit und Ausblick **9**

4 Anhang **10**
Glossar I
Abbildungsverzeichnis II
Tabellenverzeichnis III
Quellcodeverzeichnis IV

1 Einleitung

Hier könnte eine Einleitung stehen *Hervorgehoben*

1.1 Überschrift abc

Abcdef ghijklmno pqrstuvw xyz äöü. . . Duale Hochschule Baden Württemberg (DHBW)

2 Use Cases

As MapReduce is an important paradigm for solving big and costly problems today, there are really many use cases. Almost every problem that can be divided into small pieces and processed independent can be distributed and calculated in parallel by following the MapReduce paradigm. Not only scientific calculation, but also today's companies are trying to solve their problems by using Map Reduce.

The following shows an exemplary list of problems that can be efficiently solved by MapReduce:

- Factorization of (big) Integers [4]
- Matrix factorization [6]
- Multiplying terabit integers [7]
- Fourier transformations
- Genetic analysis [2]
- Search engine indexing
- Big sort problems [1]
- Users' behavior analytics
- Group text documents into topically related groups
- Extract Transform Load (ETL) and Data Mining
- Friend finder
- Defending email spam (Yahoo Mail) [5]
- Calculation of navigation / shortest path problems [1]
- Word count [1]
- Weather prediction

Some of these use cases are more scientific, others are for commercial use. Especially while data warehouses are rising, there is a need for calculating big data. When data is filled into the data ware house, there are often calculations to be done like calculating aggregations. This can be efficiently done by using Map Reduce. When data is already filled inside the data ware house, there are several techniques for data mining. As many parts of data mining can be split into pieces and calculated separately, it is very efficient to take Map Reduce algorithms to do so.

As MapReduce processing is computed by large clusters, nowadays with many cloud providers it is more a matter of being interested in the results than having the money to buy an own cluster. Almost all providers of cloud computing are having resources or hardware optimized solutions for doing MapReduce. E.g. Amazons Elastic Map Reduce (EMR) is between $0.015 and$0.50 per hour (Quelle: http://aws.amazon.com/de/elasticmapreduce/pricing/), which isn't really expensive - especially if it is for commercial use.

Genetic analysis is an upcoming trend. Today the costs for a complete genetic analysis are below $1000. By doing this analysis on public cloud infrastructure and doing the analysis with the Map Reduce paradigm, these costs could be significantly reduced. Maybe in some years, it will be a normal thing to analyses the genomes of a person? - The technic is ready to do so. The bigger question is: Are the people willing to do that?

Source: http://www.medcitynews.com [3]

Figure 2.1: Picture of a genome

Factorization of big integers is a very interesting but also a dangerous use case. Many security and cryptographic methods are based on the problem that it is not possible to efficiently factorize big integers. Map Reduce is a way to solve some factorization problems. But if someone could find a really efficient way (much more efficient than Map Reduce), this would lead to new challenges in cryptography.

Search engine companies like Google do have a big need for indexing huge amount of data to give their users the best and most relevant search results in the shortest possible time. It is hardly surprising that Google as the market leader and inventor of MapReduce uses MapReduce to do that job while it is highly parallelizable. Google also uses the same techniques for indexing Emails in Gmail.

Friend finders were a revolutionary new thing for social networks. Companies. With Facebook's friend finder leading the way big social networks implemented better search engines for people who are searching for their friends. A large amount of data has to be analyzed to give good suggestions for possible friends. Indexing profiles and searching for people can be highly parallelized. Facebook implemented its friend finder by using the MapReduce pattern.

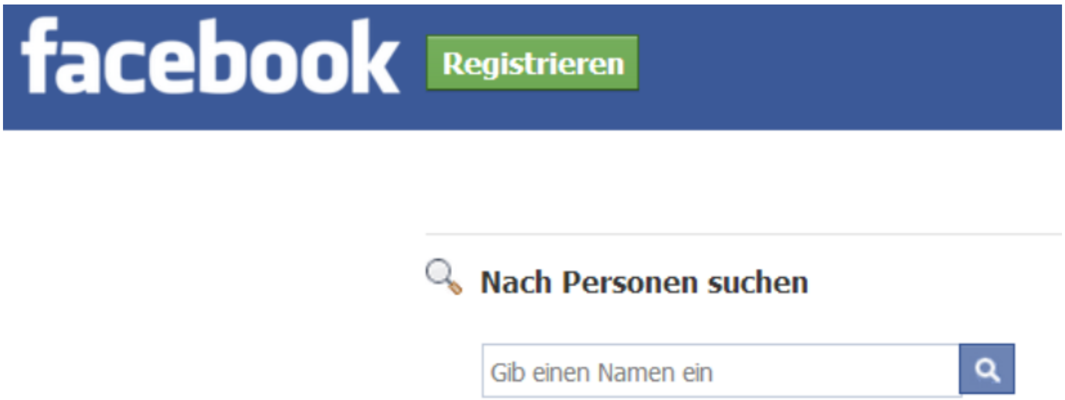

Figure 2.2: Facebooks friend finder

3 Fazit und Ausblick

4 Anhang

Glossar

DHBW Duale Hochschule Baden Württemberg

ETL Extract Transform Load

List of Figures

2.1 Picture of a genome 7
2.2 Facebooks friend finder 8

List of Tables

Quellcodeverzeichnis

Bibliography

[1] AARON: *MapReduce Algorithms (Lecture at University of Washington)* *http://courses.cs.washington.edu/courses/cse490h/08au/lectures/algorithms.pdf, Seen at 17th of March 2014*. October 2008

[2] AARON MCKENNA & MATTHEW HANNA & ERIC BANKS & ANDREY SIVACHENKO & KRISTIAN CIBULSKIS & ANDREW KERNYTSKY & KIRAN GARIMELLA & DAVID ALTSHULER & STACEY GABRIEL & MARK DALY & AND MARK A. DEPRISTO: The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data
http://genome.cshlp.org/content/20/9/1297.long, Seen at 15th of March 2014 / The Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA;
Center for Human Genetic Research, Massachusetts General Hospital, Richard B. Simches Research Center, Boston, Massachusetts 02114, USA. 2009. – Forschungsbericht

[3] DEANNA POGORELC: *medcitynews.com* *http://medcitynews.com/wp-content/uploads/dna-588x284.jpg, Seen at 12th of March 2014*. December 2012

[4] JAVIER TORDABLE: *Map Reduce for Integer Factorization* *https://code.google.com/p/mapreduce-integer-factorization/, Seen at 12th of March 2014*. December 2009

[5] MIKE KIRKWOOD: *Map Reduce your Inbox: Yahoo Mail is Fighting Spam with Big Data* *http://readwrite.com/2010/05/24/map-reduce-yahoo-mail#awesm=~ozyRktADsuJsXT, Seen at 19th of March 2014*. May 2010

[6] SEBASTIAN SCHELTER & CHRISTOPH BODEN & MARTIN SCHENCK & ALEXANDER ALEXANDROV & VOLKER MARKL: Distributed Matrix Factorization with MapReduce using a series of Broadcast-Joins https://code.google.com/p/mapreduce-integer-factorization/, seen at 11th of March 2014 / Technische Universität Berlin, Germany. 2013. – Forschungsbericht

[7] TSZ-WO SZE: Schönhage-Strassen Algorithm with MapReduce for Multiplying Terabit Integers https://people.apache.org/~szetszwo/ssmr20110430.pdf, seen at 12th of March 2014 / Yahoo! Cloud Platform. 701 First Avenue, Sunnyvale, CA 94089, USA, April 2011. – Forschungsbericht