

Neo4j

Übungsprojekt: noSQL



1. Dezember 2013

Valentin Bürgler, Fabian Meyer, Larissa Schuler, Maurin Thalmann

DBS

Titel, Namen, Modul, Semester

1. Inhalt

[1 Einführung 2](#_Toc500321018)

[1.1 Was ist der Kontext, warum ist das Projekt relevant, und worum geht es? 2](#_Toc500321019)

[2 Datenmanagement 2](#_Toc500321020)

[2.1 Um welche Datenbanktechnologie handelt es sich? 2](#_Toc500321021)

[2.2 Welche Anwendung (Use Case) unterstützt ihre Datenbank? 2](#_Toc500321022)

[2.3 Welche Daten werden migriert / eingefügt, und wie genau? 2](#_Toc500321023)

[2.4 Wie interagiert der Benutzer mit der Datenbank? 2](#_Toc500321024)

[3 Datenmodellierung 2](#_Toc500321025)

[3.1 Welches Datenmodell (ER) liegt ihrem Projekt zugrunde? 2](#_Toc500321026)

[3.2 Wie wird ihr Datenmodell in Ihrer Datenbank in ein Schema übersetzt? 2](#_Toc500321027)

[4 Datenbanksprachen 2](#_Toc500321028)

[4.1 Wie werden Daten anhand einer Query abgefragt? 2](#_Toc500321029)

[5 Konsistenzsicherung4 2](#_Toc500321030)

[5.1 Wie wird die Datensicherheit gewährleistet? 2](#_Toc500321031)

[5.2 Wie können Transaktionen parallel / konkurrierend verarbeitet werden? 2](#_Toc500321032)

[6 Systemarchitektur 2](#_Toc500321033)

[6.1 Wie ist der Server aufgebaut und wie wurde er installiert? 2](#_Toc500321034)

[6.2 Wie kann die Effizienz von Datenanfragen optimiert werden? 2](#_Toc500321035)

[7 Vergleich mit relationalen Datenbanken 2](#_Toc500321036)

[7.1 Vergleichen Sie ihre NoSQL-Technologie mit SQL-Datenbanken. 2](#_Toc500321037)

[8 Schlussfolgerungen 2](#_Toc500321038)

[8.1 Was haben Sie erreicht, und welche Erkenntnisse haben sie dabei gewonnen? 2](#_Toc500321039)

[8.2 Wie beurteilt ihre Gruppe die gewählte Datenbanktechnologie, und was sind Vor- und 2](#_Toc500321040)

[8.3 Nachteile? 2](#_Toc500321041)

# Einführung

## Was ist der Kontext, warum ist das Projekt relevant, und worum geht es?

# Datenmanagement

## Um welche Datenbanktechnologie handelt es sich?

Wir benutzen für dieses Projekt die Datenbank Neo4j. Diese ist eine Open-Source-Graphdatenbank, welche mit Java implementiert wurde. Sie erschien 2007 und ist im Vergleich mit anderen Datenbanken auf dem Platz 21 und die beste Graphdatenbank ([www.db-engines.com](http://www.db-engines.com); Stand: 06.12.2017).

## Welche Anwendung (Use Case) unterstützt ihre Datenbank?

Zu den Use Cases der Neo4j-Datenbank (und allgemein von Graphdatenbanken) zählt zum Beispiel Betrugsentdeckung (fraud detection). Dabei müssen Datenbeziehungen in Echtzeit analysiert werden können, um Betrüger frühzeitig erkennen zu können.

Auch sogenannte Knowledge Graphs können dadurch dargestellt werden. Ein Knowledge Graph kann Antworten auf bestimmte Fragen liefern, indem alle möglichen Daten gesammelt und dazu Beziehungen hergestellt werden.

Network&IT:

Real-Time Recommendation Engines:

Master Data Management:

Social Network:

Identity&Access Management

https://neo4j.com/use-cases/

Use cases include matchmaking, network management, software analytics, scientific research, routing, organizational and project management, recommendations, social networks, and more.

## Welche Daten werden migriert / eingefügt, und wie genau?

Wenn man eine SQL-Datenbank in Neo4J importieren will, muss man zuerst alle Tabellen als CSV-Dateien exportieren. In Neo4J kann man die Dateien dann mit dem Cypher-Befehl:

LOAD CSV FROM 'file:/assistenten.csv' AS assistenten CREATE (:Assistenten { PersNr: assistenten[0], Name: assistenten[1], Fachgebiet: assistenten[2], Boss: assistenten[3] })

Importieren. Wichtig ist zu beachten, dass Neo4J über einen eigenen Import-Folder verfügt und durch file:/ automatisch dort nach der Datei sucht. Weiterhin muss man beim Importieren auch noch alle Informationen bezüglich der Benennung der Spalten der jeweiligen Tabelle angeben.

Beziehungen stellt man in Neo4J mit dem folgenden Cypher-Befehl her:

MATCH (a:Assistenten), (p:Professoren) WHERE a.Boss = p.PersNr CREATE (a)-[:Ist\_angestellt\_von]->(p)

**Dies stellt bei allen Assistenten und Professoren eine Beziehung mit dem Namen „Ist angestellt von“ her, bei denen der Boss von a (Assistenten) gleich der PersNr von p (Professoren) ist.**

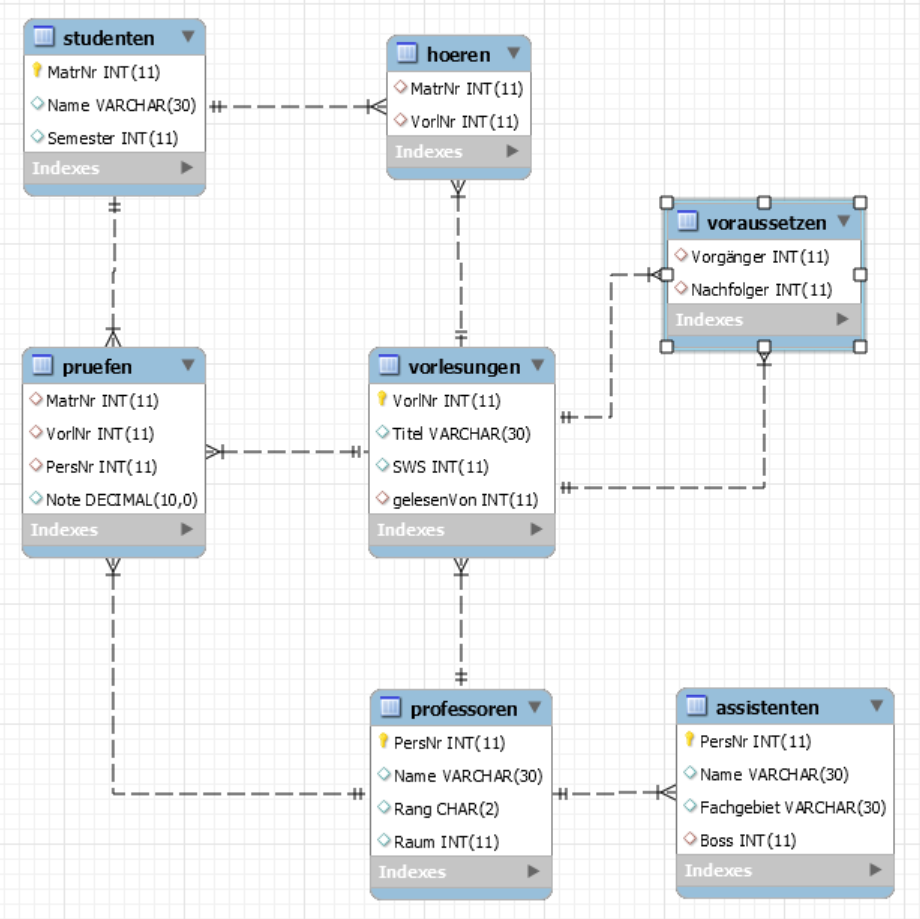
## Wie interagiert der Benutzer mit der Datenbank?

Als Benutzer nutzt man die Abfragesprache Cypher.

# Datenmodellierung

## Welches Datenmodell (ER) liegt ihrem Projekt zugrunde?

Das SQL-Datenmodell sieht folgendermassen aus:



## Wie wird ihr Datenmodell in Ihrer Datenbank in ein Schema übersetzt?

Bei Neo4j wird das Datenmodell durch Knoten, Kanten und deren Beziehungen zu einander dargestellt.

# Datenbanksprachen

## Wie werden Daten anhand einer Query abgefragt?

# Konsistenzsicherung4

## Wie wird die Datensicherheit gewährleistet?

## Wie können Transaktionen parallel / konkurrierend verarbeitet werden?

# Systemarchitektur

## Wie ist der Server aufgebaut und wie wurde er installiert?

## Wie kann die Effizienz von Datenanfragen optimiert werden?

# Vergleich mit relationalen Datenbanken

## Vergleichen Sie ihre NoSQL-Technologie mit SQL-Datenbanken.

Neo4j kann viel besser mit einer grossen Menge an Daten und vielen Beziehungen zwischen den Daten umgehen und somit leistet sie eine bessere Performance als SQL-Datenbanken.

# Schlussfolgerungen

## Was haben Sie erreicht, und welche Erkenntnisse haben sie dabei gewonnen?

## Wie beurteilt ihre Gruppe die gewählte Datenbanktechnologie, und was sind Vor- und

## Nachteile?