



**Hochschule  
Bonn-Rhein-Sieg**  
University of Applied Sciences

**Schemalose Datenbank**

# **MongoDB als Datenbank für Document Stores**

**Falls erforderlich: Zur Erlangung des akademischen Grades eines  
Master of Science  
- Studiengang MCS -**

Fachbereich Informatik  
Referent: Prof. Dr. Knolle  
Falls erforderlich: Korreferent: B. Mager

eingereicht von:  
Franz-Dominik Dahmann  
Matr.-Nr. 9019745  
Münchhausenweg 30  
53340 Meckenheim

Sankt Augustin, den 06.11.2016

## Zusammenfassung

Moderne Datenbanken sehen sich mit diversen Problemen konfrontiert. Es sollen extrem große Datenmengen verarbeitet werden und die Antwortzeiten sollen niedrig gehalten werden. Ferner soll es für sehr viele Nutzer gleichzeitig möglich sein Abfragen zu stellen. Aus diesen Gründen ist es in modernen Datenbanken notwendig ist, Datenstrukturen möglichst flexibel und performant zu speichern. Diese Probleme sollen mit verschiedenen Ansätzen aus dem Umfeld der NoSQL Datenbanksysteme gelöst werden. Die dokumentenbasierte Speicherung versucht für Probleme, welche sich meist auf große Mengen mit heterogenen Daten beziehen, einen Lösungsansatz zu finden. In diesem Umfeld ist die Funktionsweise der MongoDB äußerst populär, da sie auf hohe Leistung, große Datenmengen, hohe Flexibilität und einfache Skalierbarkeit ausgelegt ist. Die Grundlage für die der MongoDB zugrunde liegenden dokumentenbasierte Speicherung bieten die Key-Value Stores, welche unformatierte Werte unter einem Key persistieren. Document Stores vereinen den Vorteil der Schemafreiheit der Schlüssel-Wert Zuordnungen mit der Möglichkeit zur Strukturierung von Daten. Dadurch sind Datenbanken wie MongoDB in der Lage komplexere JSON ähnliche Datentypen zu verarbeiten und basierend auf ihren Attributen zu indexieren. Die hier umrissenen theoretischen Grundlagen werden detaillierter beschrieben und in dem NoSQL Kontext bewertet.

Um die theoretischen Grundlagen zu verdeutlichen werden mithilfe eines Anwendungsszenarios die Vorteile von MongoDB verdeutlicht. Grundlage für das Anwendungsszenario ist das logische Schema einer relationalen Datenbank, in der sowohl Abschlussarbeitsthemen, als auch alle beteiligten Personen, wie beispielsweise Studenten oder Prüfer, sowie Organisationen, gespeichert sind.

Ziel des Anwendungsszenarios ist es, dieses Schema mit Hilfe von MongoDB umzusetzen. Ferner soll die Eignung von MongoDB in diesem Kontext analysiert und bewertet werden. Damit MongoDB in diesem Kontext bewertet werden kann, werden unter anderem Performancetests an der Datenbank durchgeführt. Weiter soll eine Web-Anwendung implementiert werden, welche die grundsätzlichen Datenbankoperationen an der MongoDB veranschaulicht.

[Edl11] [Sad13] [Red12] [Han11] [Tud11] [And16b] [Dan16b] [And16a] [Dan16a] [Boi12]  
[Par13] [Abr13] [Far14] [Sha15] [How14]

# Inhaltsverzeichnis

<b>Zusammenfassung</b>	<b>II</b>
<b>1 Ausgewählte Persistenzmodelle</b>	<b>1</b>
1.1 Key-Value Stores . . . . .	1
1.1.1 Einleitung . . . . .	1
1.1.2 Eigenschaften . . . . .	1
1.1.3 Datenmodell . . . . .	1
1.1.4 Abgrenzung . . . . .	1
1.1.5 Anwendungsfälle . . . . .	1
1.1.6 Bewertung . . . . .	1
1.2 Document Stores . . . . .	1
1.2.1 Einleitung . . . . .	1
1.2.2 Eigenschaften . . . . .	1
1.2.3 Datenmodell . . . . .	2
1.2.4 Anwendungsfälle . . . . .	3
1.2.5 Bewertung . . . . .	3
<b>2 Persistenzsysteme und Big Data Frameworks</b>	<b>4</b>
2.1 MongoDB . . . . .	4
2.1.1 Einleitung . . . . .	4
2.1.2 Geschichte . . . . .	4
2.1.3 Anwendungsfälle . . . . .	4
2.1.4 Einrichtung . . . . .	4
2.1.5 Verwendung . . . . .	4
2.1.6 Grenzen . . . . .	4
2.1.7 Zukunft . . . . .	4
<b>3 Ausgewählte Anwendungsszenarien</b>	<b>5</b>
3.1 Web 2.0 - Abschlussarbeiten-DB . . . . .	5
3.2 MongoDB . . . . .	5
<b>Literaturverzeichnis</b>	<b>6</b>

# **1 Ausgewählte Persistenzmodelle**

## **1.1 Key-Value Stores**

### **1.1.1 Einleitung**

### **1.1.2 Eigenschaften**

### **1.1.3 Datenmodell**

### **1.1.4 Abgrenzung**

### **1.1.5 Anwendungsfälle**

### **1.1.6 Bewertung**

## **1.2 Document Stores**

### **1.2.1 Einleitung**

Document-Stores speichern strukturierte Daten in Datensätzen, die Dokumente genannt werden. Grundsätzlich vereinigen sie die Schemafreiheit von Key-Value-Stores mit der Möglichkeit zur Strukturierung der gespeicherten Daten. Somit lassen sich Document-Stores als Erweiterung von Key-Value Stores betrachten. Vorteil hierbei ist, dass es bei Document-Stores im Gegensatz zu Key-Value Stores möglich ist komplexere auf den Inhalt (die Attribute) der Dokumente gerichtete Abfragen zu stellen. Dies ist bei Key-Value-Stores nicht möglich.

### **1.2.2 Eigenschaften**

Die Kerneigenschaften der Document-Stores sind:

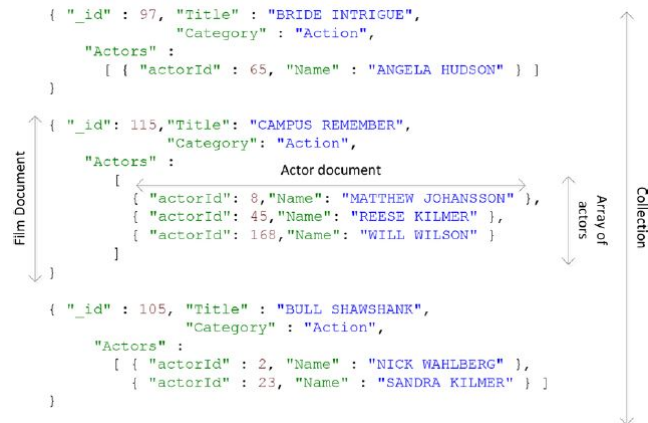


Abbildung 1.1: Beispiel eines JSON-Dokuments

- Daten in Form von Dokumenten
- Dokumente sind Schemalos
- Horizontale Verteilung

Bei der Betrachtung von Document-Stores bezieht sich das Wort "document" auf die hierarchische Struktur der abgespeicherten Daten. Die hier betrachteten Dokumente bestehen jedoch lediglich aus binären Daten oder einfachem Text. Ferner kann es sich um Semi-strukturierte Daten wie JSON oder XML handeln. Während ältere Document-Stores auf XML setzten, ist bei neueren Document-Stores das JSON ähnliche BSON üblich.

### 1.2.3 Datenmodell

Im Zentrum des Datenmodells der Document-Stores stehen die Documents, die eine hierarchische Struktur aufweisen. So kann ein Dokument auf viele weitere Dokumente verweisen. Dies ist möglich, da ein Dokument Listen enthalten kann. Die Dokumente liegen üblicherweise in JSON bzw. BSON vor.

## JSON

Die Java-Script-Object-Notation ist ein vom Menschen und Maschinen lesbarer Standard zur Beschreibung von Daten. JSON ist wie XML ein Standard zum Datenaustausch im Web. In 1.1 ist ein Beispiel für ein JSON-Dokument zu sehen. Vorteil von JSON ist, dass JSON-Dokumente einfach zu zerlegen und umzuwandeln sind. Dies verringert den Aufwand, der in der Anwendungsschicht betrieben werden muss.

## **BSON**

Unter BSON sind binär kodierte JSON-Dokumente zu verstehen. BSON erweitert das Modell von JSON um weitere Datentypen und ermöglicht effizientes kodieren und dekodieren innerhalb verschiedener Sprachen.

### **1.2.4 Anwendungsfälle**

JSON basierte Document-Stores haben vor allem in Web-basierten Anwendungen ihre Vorzüge. Hier wird häufig JSON als Datenschicht verwendet, weswegen Document-Stores hier bevorzugt werden.

XML-basierte Document-Stores finden vor allem in Content-Management-Systemen Anwendung, da sie hier ein Management Repository für XML-basierte Textdateien zur Verfügung stellen.

### **1.2.5 Bewertung**

## **2 Persistenzsysteme und Big Data Frameworks**

### **2.1 MongoDB**

#### **2.1.1 Einleitung**

#### **2.1.2 Geschichte**

#### **2.1.3 Anwendungsfälle**

#### **2.1.4 Einrichtung**

#### **2.1.5 Verwendung**

#### **2.1.6 Grenzen**

#### **2.1.7 Zukunft**



## **3 Ausgewählte Anwendungsszenarien**

### **3.1 Web 2.0 - Abschlussarbeiten-DB**

### **3.2 MongoDB**

# Literaturverzeichnis

- [Abr13] ABRAMOVA, Veronika: *NoSQL Databases: MongoDB vs Cassandra*. ACL Digital Library, 2013
- [And16a] ANDREAS, Meier: *Datenmanagement*. Springer, 2016
- [And16b] ANDREAS, Meier: *Zur Nutzung von SQL- und NoSQL-Technologien*. Springer, 2016
- [Boi12] BOICEA, Alexandru: *MongoDB vs Oracle – Database Comparison*. Researchgate, 2012
- [Dan16a] DANIEL, Fasel: *Übersicht über NoSQL-Technologien und -Datenbanken*. Springer, 2016
- [Dan16b] DANIEL, Fasel: *Was versteht man unter Big Data und NoSQL*. Springer, 2016
- [Edl11] EDLICH, Stefansub: *NoSQL: Einstieg in die Welt nichtrelationaler Web 2.0 Datenbanken*. 2. München Hanser Verlag, 2011
- [Far14] FARAJ, Azhi: *Comparative study of Relational and Nonrelational Database performances using Oracle and MongoDB Systems*. IJCET, 2014
- [Han11] HAN, Jing: *Survey on NoSQL database*. IEEE, 2011
- [How14] HOWS, David: *MongoDB Basics a quick introduction to MongoDB*. APRESS, 2014
- [Par13] PARKER, Zachary: *Comparing NoSQL MongoDB to an SQL DB*. ACL Digital Library, 2013
- [Red12] REDMOND, Eric: *Sieben Wochen, sieben Datenbanken*. O'Reilly Verlag, 2012
- [Sad13] SADALAGE, Pramod J.: *NoSQL: A Brief Guide to the Emerging World of Polyglot Persistence*. Addison-Wesley, 2013
- [Sha15] SHAKUNTALA, Gupta E.: *Practical MongoDB Architecting, Developing, and Administering MongoDB*. APRESS, 2015
- [Tud11] TUDORICA, Bogdan G.: *A comparison between several NoSQL databases with comments and notes*. IEEE, 2011