*Related Work*

1. Benchmarking Cloud Serving Systems with YCSB (Juni 2010)

Betrachtete Datenbanken:

* Cassandra
* HBase
* PNUTS
* sharded MySQL

Stärken:

* neben den „Standard“- Workloads wird zusätzlich Skalierbarkeit untersucht

Schwächen:

- Verfügbarkeit (Availability) wird nicht untersucht (Kernaspekt von Cloud Anwendungen)

- Replizierbarkeit wird nicht untersucht

- encrypted data wird nicht betrachtet

Kommentar:

* Wäre gut, Insert, Update, Read, Scan wie hier definieren

Link: [Benchmarking cloud serving systems with YCSB (acm.org)](https://dl.acm.org/doi/pdf/10.1145/1807128.1807152?casa_token=b8oZr3iNm4cAAAAA:5SGeqy9F5YWz89Hr8R-X4CGmV1OqQCK4OLYC1yAZQeuOq1srqw7ZcuCG2YbovCGY9403X1hWWJo))

1. A Study over NoSQL Performance (April 2019)

Betrachtete Datenbanken:

* Redis
* Memcached
* Voldemort
* MongoDB
* OrientDB
* Cassandra
* HBase

Stärken:

* Es werden mindestens zwei Datenbanken pro Typ (Key-value, Dokument-basierte, Column-orientierte) betrachtet

Schwächen:

* Nur eine persistente Key-Value Datenbank wird betrachtet (unserem Fall 2 Datenbanken pro Typ)

Link: [(PDF) A Study over NoSQL Performance (researchgate.net)](https://www.researchgate.net/publication/332028074_A_Study_over_NoSQL_Performance)

1. Performance Benchmarking of Key-Value Store NoSQL Databases (2018)

Betrachtete Datenbanken:

-Redis

-SSDB

Stärken:

Schwächen:

* Nur zwei Datenbanken, wobei SSDB auch Redis Client unterstützt

Kommentar:

* Es gibt vllt nix, was wir von hier nehmen/lernen können
* Erinnerung: Graphs in geeignete Auflösung darstellen

1. Which NoSQL Database? A Performance Overview (2014)

Betrachtete Datenbanken:

* Cassandra
* HBase
* MongoDB
* OrientDB
* Redis

Stärken:

* Hohe Anzahl von Records (600000) und Operationen über den Records (1000 je)

Schwächen:

* Nur eine key value Datenbank
* Nur 3 Workloads

Kommentar:

* Hoch, aber auch ausreichende Anzahl von Test-Operationen wählen
* Scalability nicht geprüft

1. Performance Evaluation of NoSQL Databases: A Case Study (2015)

Betrachtete Datenbanken:

* MongoDB
* Cassandra
* Riak

Stärken:

* Replizierbarkeit geprüft (32 Sessions/Versuche)

Schwächen:

* Nur eine persistente key-value Datenbank
* Workloads nur Read/Write
* Graphs schlecht zu lesen

Kommentar:

* Scalability mit Latency distribution zeigen (Read/Write)

1. A performance comparison of SQL and NoSQL databases (2013)

Betrachtete Datenbanken:

* MongoDB
* RavenDB
* CouchDB
* Cassandra
* Hypertable
* Couchbase
* MS SQL Express

Stärken:

* Database bucket instantiation (Datenbank starten?) auch getestet (in ms)
* Delete, key fetching auch getestet
* Scalability gemessen/gezeigt mit steigenden Anzahl von Operationen

Schwächen:

* Keine key-value Datenbank
* Tabellen mit Nummer ein bisschen schwer zu lesen

Link: [A performance comparison of SQL and NoSQL databases | IEEE Conference Publication | IEEE Xplore](https://ieeexplore.ieee.org/document/6625441)

1. Performance Study of SQL and NoSQL Solutions for Analytical Loads (2014)

Betrachtete Datenbanken:

* MongoDB
* Vertica
* Oracle

Stärken:

* Simplere Vergleichungstabelle die Datenbanken

Schwächen:

* Nur eine NoSQL Datenbank wird betrachtet im Vergleich zu zwei SQL Datenbanken

Kommentar:

* Vor-Filtering für bessere Leistung bei Get/Query?

Generelle Schwächen in den Papern:

* Wie beeinflusst die Replizierbarkeit die Aktualität von Daten?
* Sind die Daten, die wir reingeschrieben haben auch wirklich so in der DB abgespeichert?

Vergleich Benchmarks:

* YCSB: ([brianfrankcooper/YCSB: Yahoo! Cloud Serving Benchmark (github.com)](https://github.com/brianfrankcooper/YCSB))

No data scalability

Workload Generation

🡪NoSQL

* CloudSite:

Partial data scalablility

Hadoop, GraphLab, NoSQL

🡪 Architectures

* BigDataBench:

Total data scalability

BDGS nicht anwendbar für unsere Hash Generation