

Key-value databázové systémy

Key-Value Database Management Systems

Bc. Jan Jedlička

Vedoucí: prof. Ing. Michal Krátký, Ph.D.

FEI, VŠB-TUO

2024



- Studium problematiky key-value databázových systémů (KDBS [1])
- Návrh a implementace testovací prostředí pro porovnání KDBS s ostatními DBS (YCSB [2])
- Otestování vybraných KDBS a vyhodnocení výsledků experimentů (Redis [3], Aerospike [4], Memcached [5], Riak KV [6])

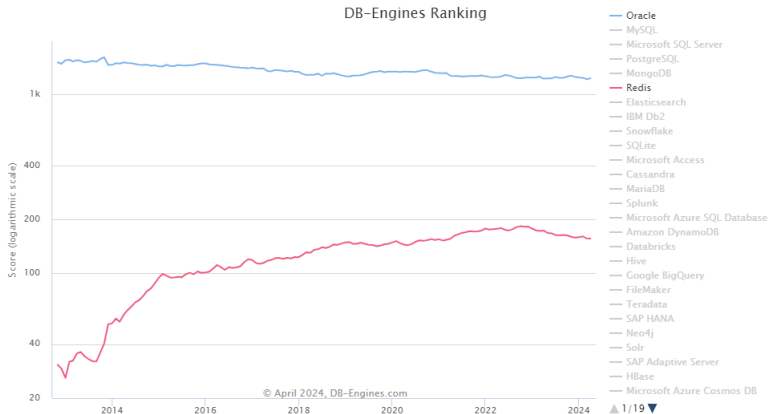
Not Only SQL (NoSQL)

- Data nezávislá na schématu (schema-less)
- Zpracování velkých objemů vzájemně nesouvisejících nebo rychle se měnících dat [7]
- Vyžadujeme výkon a dostupnost oproti silné konzistenci
- Distribuované (sdílené) úložiště
- Horizontální škálovatelnost

Not Only SQL (NoSQL)

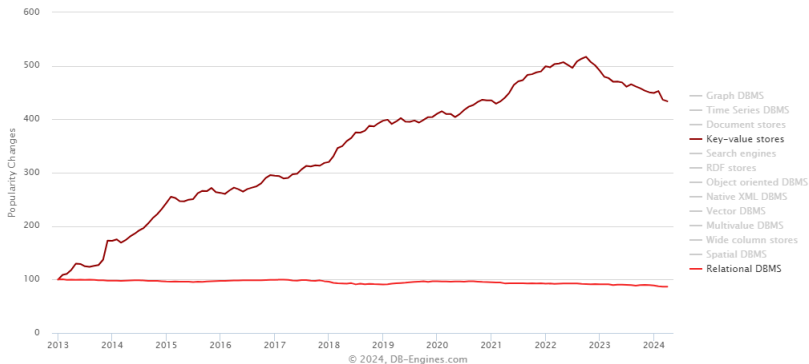
- Klíč-hodnota DBS (Redis [3], Aerospike [4])
- Dokumentové DBS (MongoDB [8], CouchDB [9])
- Sloupcové DBS (Apache Cassandra [10], Apache HBase [11])
- Grafové DBS (Neo4J [12], Azure Cosmos DB [13])

RDBS Oracle vs KDBS Redis



Obrázek: Graf hodnot popularity RDBS Oracle a KDBS Redis [14]

RDBS vs KDBS



Obrázek: Graf změn hodnot popularity RDBS a KDBS [15]

- Databáze párů klíč-hodnota libovolného formátu
- Klíč je unikátní identifikátor umožňující rychlý přístup k hodnotám
- Základní operace (`get(k)`, `insert(k,h)`, `delete(k)`, `update(k,h)`) [16]

```
1 {  
2   "jmeno": "Jan Novak",  
3   "vek": 35,  
4   "adresa": null,  
5   "kontakt": [  
6     {"email": "jan.novak.01@gmail.com"},  
7     {"telefon": "406-792-448"}  
8   ]  
9 }
```

Obrázek: Hodnota (JSON dokument) uložená pro klíč "userid_135"

TPC-H - SQL Test Q1

```
1  --Q1
2  select
3      l_returnflag,
4      l_linestatus,
5      sum(l_quantity) as sum_qty,
6      sum(l_extprice) as sum_base_price,
7      sum(l_extprice * (1 - l_discount)) as sum_disc_price,
8      sum(l_extprice * (1 - l_discount) * (1 + l_tax)) as sum_charge,
9      avg(l_quantity) as avg_qty,
10     avg(l_extprice) as avg_price,
11     avg(l_discount) as avg_disc,
12     count(*) as count_order
13 from lineitem
14 where l_shipdate <= date '1998-12-01'
15 group by l_returnflag, l_linestatus
16 order by l_returnflag, l_linestatus;
```

Obrázek: TPC-H, SQL Test Q1 [17, 18]

YCSB - operace vložení

```
1 {  
2   "operation": "insert",  
3   "table": "usertable",  
4   "key": "user1474",  
5   "values": {  
6     "field0": "value0",  
7     "field1": "value1",  
8     "field2": "value2"  
9   }  
10 }
```

Obrázek: YCSB - příklad definice operace vložení záznamu

- Java JDK 8 (1.8) [19]
- Nastavení %JAVA_HOME% systémové proměnné [20] (OS Windows)
- Docker [21]
- YCSB 0.17 [22]
- Apache Maven 3 [23]

Zprovoznění testů - Docker

```
PS E:\ycsb-0.17.0\bin> docker pull redis:latest
latest: Pulling from library/redis
Digest: sha256:f14f42fab123example456c7e824b9
Status: Image is up to date for redis:latest
docker.io/library/redis:latest

PS E:\ycsb-0.17.0\bin> docker run --name my-redis -p 6379:6379 -d redis:latest
aba5fab123example4568aa97

PS E:\ycsb-0.17.0\bin> docker stop aba5fab123example4568aa97
aba5fab123example4568aa97

PS E:\ycsb-0.17.0\bin> docker rm aba5fab123example4568aa97
aba5fab123example4568aa97
```

Obrázek: Docker - příkazy pro stažení, spuštění, zastavení a odstranění DBS Redis

Zprovoznění testů - YCSB Load (1/2)

```
PS E:\ycsb-0.17.0\bin> .\ycsb load redis -P ..\workloads\workloada
      -p redis.host=127.0.0.1 -p redis.port=6379 -p recordcount=10000
      -p threadcount=4

YCSB Client 0.17.0
Loading workload...
Starting test.
[OVERALL], RunTime(ms), 4211
[OVERALL], Throughput(ops/sec), 2374.7328425552128
[TOTAL_GCS_PS_Scavenge], Count, 4
[TOTAL_GC_TIME_PS_Scavenge], Time(ms), 6
[TOTAL_GC_TIME_%_PS_Scavenge], Time(%), 0.14248397055331277
[TOTAL_GCS_PS_MarkSweep], Count, 0
[TOTAL_GC_TIME_PS_MarkSweep], Time(ms), 0
[TOTAL_GC_TIME_%_PS_MarkSweep], Time(%), 0.0
[TOTAL_GCs], Count, 4
```

Obrázek: YCSB Redis, příkaz pro vložení dat (load) - 1/2

Zprovoznění testů - YCSB Load (2/2)

```
[TOTAL_GC_TIME], Time(ms), 6
[TOTAL_GC_TIME_%], Time(%), 0.14248397055331277
[CLEANUP], Operations, 4
[CLEANUP], AverageLatency(us), 333.25
[CLEANUP], MinLatency(us), 80
[CLEANUP], MaxLatency(us), 1060
[CLEANUP], 95thPercentileLatency(us), 1060
[CLEANUP], 99thPercentileLatency(us), 1060
[INSERT], Operations, 10000
[INSERT], AverageLatency(us), 1647.61
[INSERT], MinLatency(us), 904
[INSERT], MaxLatency(us), 13895
[INSERT], 95thPercentileLatency(us), 2481
[INSERT], 99thPercentileLatency(us), 3389
[INSERT], Return=OK, 10000
```

Obrázek: YCSB Redis, příkaz pro vložení dat (load) - 2/2

Zprovoznění testů - YCSB Run (1/2)

```
PS E:\ycsb-0.17.0\bin> .\ycsb run redis -P ..\workloads\workloada
    -p redis.host=127.0.0.1 -p redis.port=6379 -p operationcount=10000
    -p threadcount=4
YCSB Client 0.17.0
Loading workload...
Starting test.
[OVERALL], RunTime(ms), 2560
[OVERALL], Throughput(ops/sec), 3906.25
[TOTAL_GCS_PS_Scavenge], Count, 1
[TOTAL_GC_TIME_PS_Scavenge], Time(ms), 1
[TOTAL_GC_TIME_%_PS_Scavenge], Time(%), 0.0390625
[TOTAL_GCS_PS_MarkSweep], Count, 0
[TOTAL_GC_TIME_PS_MarkSweep], Time(ms), 0
[TOTAL_GC_TIME_%_PS_MarkSweep], Time(%), 0.0
[TOTAL_GCs], Count, 1
[TOTAL_GC_TIME], Time(ms), 1
[TOTAL_GC_TIME_%], Time(%), 0.0390625
```

Obrázek: YCSB Redis, příkaz pro spuštění testů (run) - 1/2

Zprovoznění testů - YCSB Run (2/2)

```
[READ], Operations, 5157
[READ], AverageLatency(us), 998.8885010665116
[READ], MinLatency(us), 389
[READ], MaxLatency(us), 17695
[READ], 95thPercentileLatency(us), 1692
[READ], 99thPercentileLatency(us), 3479
[READ], Return=OK, 5157
[CLEANUP], Operations, 4
[CLEANUP], AverageLatency(us), 221.75
[CLEANUP], MinLatency(us), 60
[CLEANUP], MaxLatency(us), 638
[CLEANUP], 95thPercentileLatency(us), 638
[CLEANUP], 99thPercentileLatency(us), 638
[UPDATE], Operations, 4843
[UPDATE], AverageLatency(us), 985.7646087136072
[UPDATE], MinLatency(us), 364
[UPDATE], MaxLatency(us), 17567
[UPDATE], 95thPercentileLatency(us), 1744
[UPDATE], 99thPercentileLatency(us), 3675
[UPDATE], Return=OK, 4843
```

Obrázek: YCSB Redis, příkaz pro spuštění testů (run) - 2/2

- Počet záznamů v databázi (recordcount)
- Počet testovaných operací (operationcount)
- Název DBS
- Workload (A-C)
- Případná konzistence (Riak KV měl chybně nastavenou silnou konzistenci)

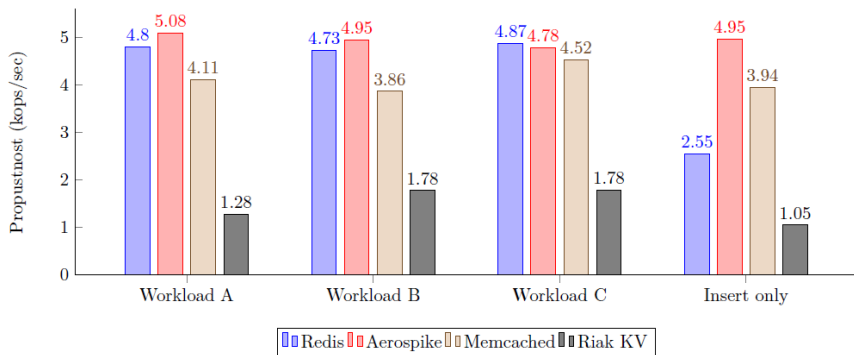
Propustnost	1.	2.	3.	4.
Průměr všech operací	Aerospike 4,9 kops/s	Redis 4,2 kops/s	Memcached 4,1 kops/s	Riak KV 1,5 kops/s
Vkládání	Aerospike	Memcached	Redis	Riak KV
Aktualizace	Aerospike	Redis	Memcached	Riak KV
Čtení	Redis	Aerospike	Memcached	Riak KV

Na třetím řádku je uvedena průměrná propustnost všech testů (tisíce operací za sekundu) pro výše vypsany KDBS

Tabulka: Porovnání propustnosti výsledků testů

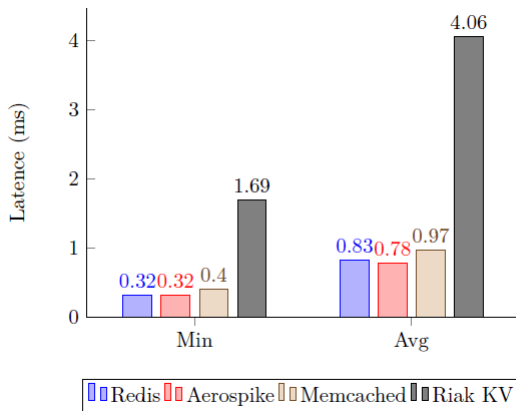
- Po překonfigurování KDBS Riak KV na případnou konzistenci (namísto silné konzistence) dosahoval systém průměrné propustnosti

Vyhodnocení testů - Propustnost (kops/sec)



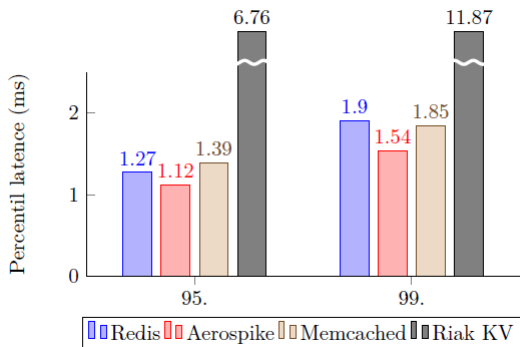
Obrázek: Workload A, B, C + Insert only - Propustnost (kops/sec)

Vyhodnocení testů - Workload A (50% čtení, 50% aktualizace)



Obrázek: Workload A - Min Avg Latence (ms)

Vyhodnocení testů - Workload A (50% čtení, 50% aktualizace)



Obrázek: Workload A - Percentil latence (ms)

Děkuji za pozornost



What Is a Key-Value Database? [online]. 2022. [cit. 2022-11-18]. Dostupné z: <https://aws.amazon.com/nosql/key-value/>.



Yahoo! Cloud Serving Benchmark (YCSB) [online]. 2019. [cit. 2024-04-22]. Dostupné z: <https://github.com/brianfrankcooper/YCSB>.



Redis [online]. 2022. [cit. 2022-11-18]. Dostupné z: <https://redis.io/>.



Aerospike [online]. 2022. [cit. 2022-11-20]. Dostupné z: <https://aerospike.com/>.



What is Memcached? [online]. 2024. [cit. 2024-04-17]. Dostupné z: <https://memcached.org/>.



Riak KV [online]. 2022. [cit. 2022-11-21]. Dostupné z: <https://riak.com/products/riak-kv/index.html>.



Databáze NoSQL [online]. 2024. [cit. 2024-08-03]. Dostupné z: <https://azure.microsoft.com/cs-cz/resources/cloud-computing-dictionary/what-is-nosql-database>.



MongoDB [online]. 2023. [cit. 2023-01-28]. Dostupné z: <https://www.mongodb.com/>.



CouchDB [online]. 2024. [cit. 2024-06-10]. Dostupné z: <https://couchdb.apache.org/>.



Cassandra [online]. 2023. [cit. 2023-01-28]. Dostupné z: https://cassandra.apache.org/_/index.html.



Welcome to Apache HBase [online]. 2024. [cit. 2024-06-10]. Dostupné z: <https://hbase.apache.org/>.



Neo4j [online]. 2024. [cit. 2024-06-10]. Dostupné z: <https://neo4j.com/>.



Azure Cosmos DB [online]. 2024. [cit. 2024-04-22]. Dostupné z: <https://azure.microsoft.com/cs-cz/products/cosmos-db>.



DB-Engines Ranking - Trend Popularity [online]. 2024. [cit. 2024-04-23]. Dostupné z: https://db-engines.com/en/ranking_trend.



DBMS popularity broken down by database model [online]. 2024. [cit. 2024-04-23]. Dostupné z: https://db-engines.com/en/ranking_categories.



BAČA, Radim. *Nerelační distribuované databáze* [online]. 2024. [cit. 2024-08-03]. Dostupné z: <https://db.cs.vsb.cz/Download.ashx?id=6>.



TPC-H Benchmark - Test SQL [online]. 2024. [cit. 2024-04-27].

Dostupné z:

https://docs.starrocks.io/docs/benchmarking/TPC-H_Benchmarking/.



TPC-H Vesion 2 and Version 3 [online]. 2024. [cit. 2024-04-27].

Dostupné z: <https://www.tpc.org/tpch/>.



Oracle - Java Downloads [online]. 2024. [cit. 2024-04-25]. Dostupné

z: <https://www.oracle.com/java/technologies/downloads/>.



Windows Environment Variables [online]. 2024. [cit. 2024-04-26].

Dostupné z: <https://ss64.com/nt/syntax-variables.html>.



Docker Builds: Now Lightning Fast [online]. 2024. [cit. 2024-03-19].

Dostupné z: <https://www.docker.com/>.



Download the latest release of YCSB [online]. 2024. [cit. 2024-04-25]. Dostupné z:

<https://github.com/brianfrankcooper/YCSB/releases/download/0.17.0/ycsb-0.17.0.tar.gz>.



Apache Maven Project [online]. 2024. [cit. 2024-04-25]. Dostupné z: <https://maven.apache.org/download.cgi>.