

Visoko skalabilen NewSQL sistem za upravljanje s podatkovnimi bazami CockroachDB

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko

Avtor: Matjaž Mav

Mentor: izr. prof. dr. Matjaž Kukar

Ljubljana, 2018

Vsebina

- Kaj je NewSQL
- Podatkovna baza CockroachDB
- Izvedba primerjalne analize
- Rezultati in ugotovitve

Kaj je NewSQL?

- SQL
- ACID transakcije
- skalabilnost
- visoka razpoložljivost
- porazdeljena okolja

Podatkovna baza CockroachDB

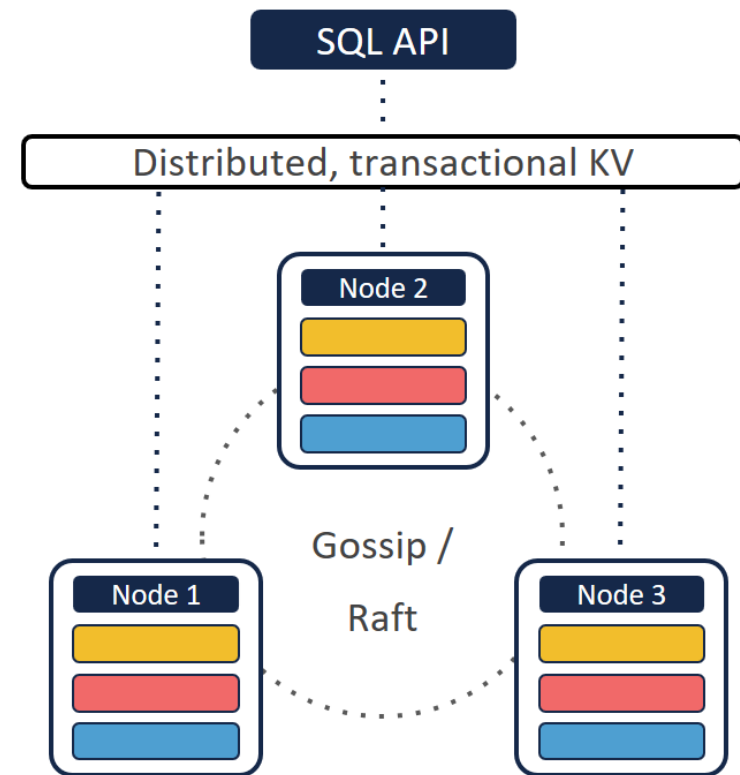
- NewSQL
- kompatibilna s PostgreSQL
- odprtokodna
- dostopna komurkoli
- enostavna za uporabo
- aktivna skupnost



Slika 1: Celostna grafična podoba podatkovne baze CockroachDB
(vir: <https://github.com/cockroachdb/cockroach>)

Arhitektura podatkovne baze CockroachDB

- plast SQL
- transakcijska plast
- porazdelitvena plast
- replikacijska plast
- shranjevalna plast



Slika 2: Arhitekturni pregled

(vir: http://cs.ulb.ac.be/public/media/teaching/cockroachdb_2017.pdf)

Postavitev testnega okolja

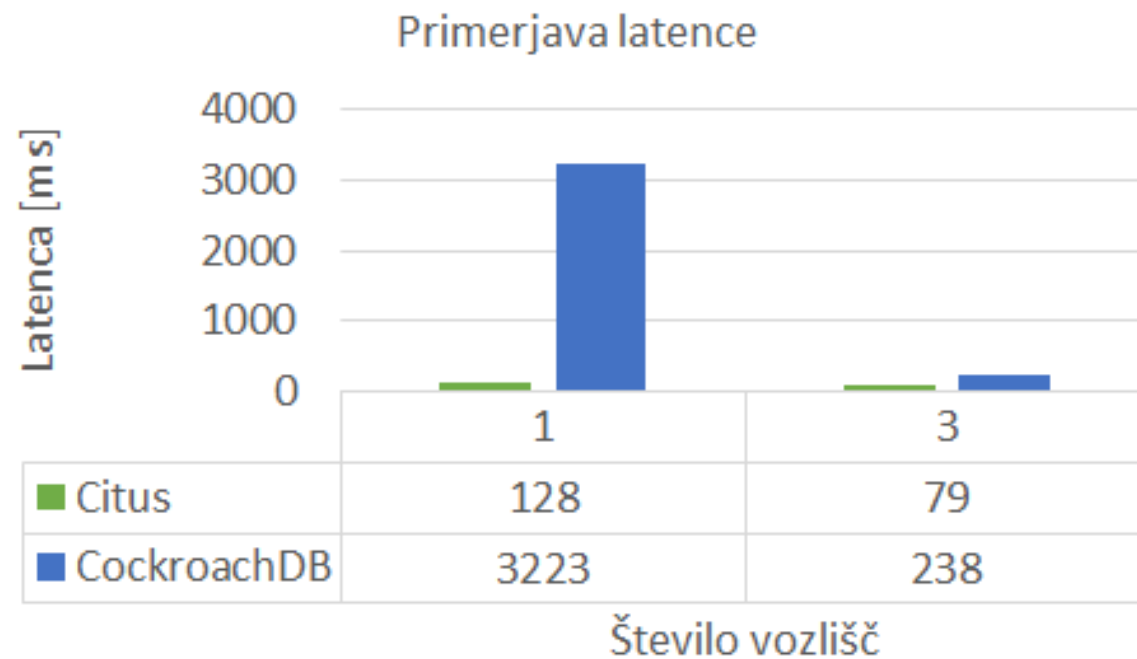
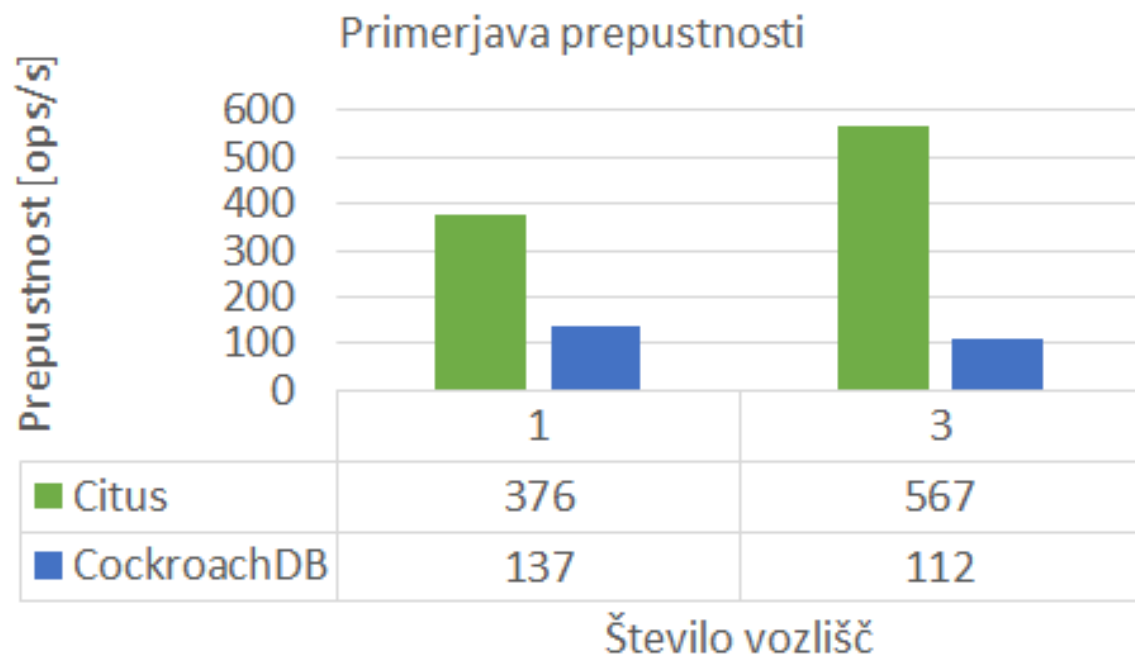
- štirje računalniki
- gigabitno omrežje
- Ubuntu Server 16.04 LTS
- Docker 18.03.0-ce
- CockroachDB 2.0.1
- PostgreSQL 10.3 + Citus 7.3.0

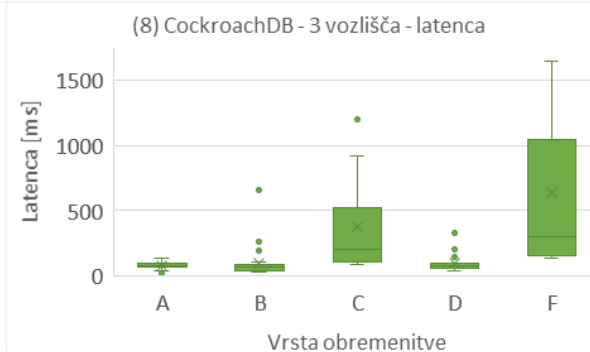
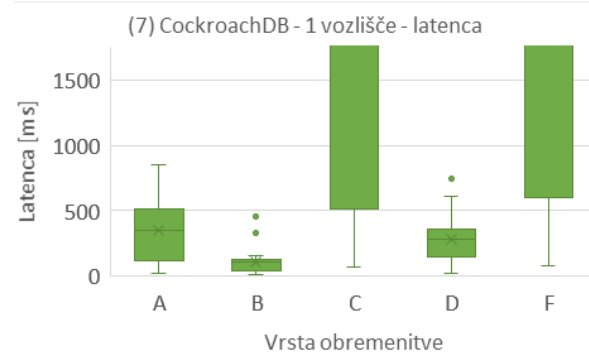
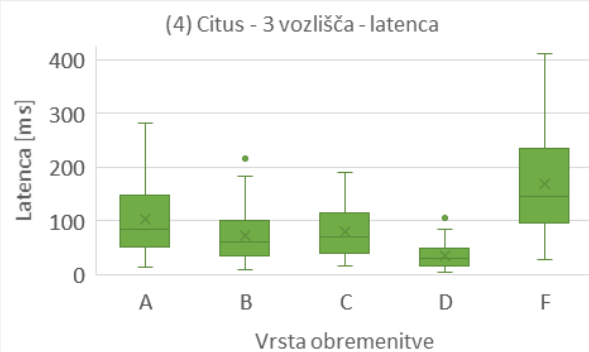
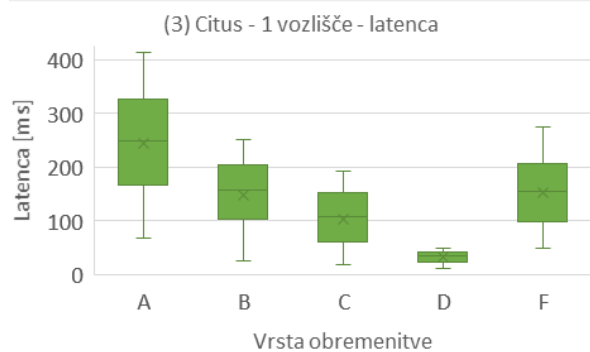
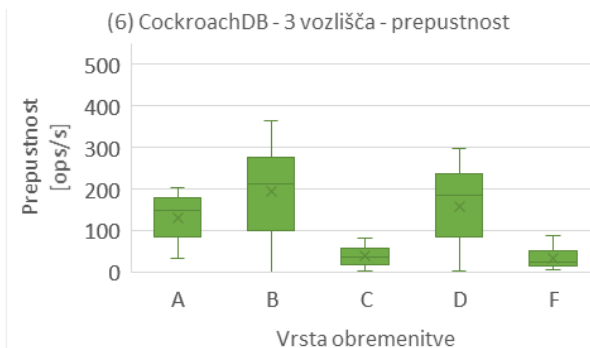
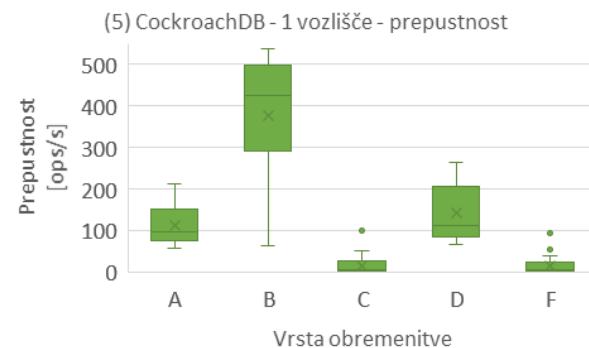
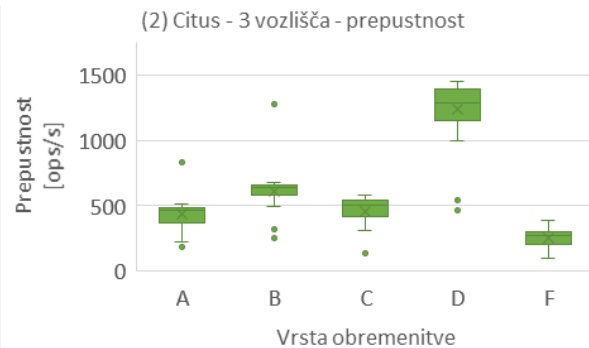
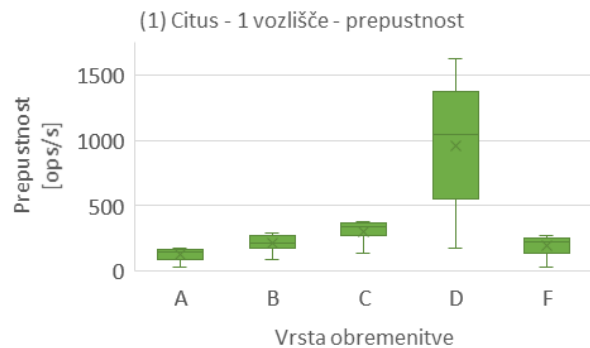
Izvedba primerjalne analize zmogljivosti

- orodje YCSB
- priprava podatkov
- parametri
- avtomatizacija testiranja
- github.com/matjazmav/diploma-ycsb

```
CREATE TABLE usertable (  
    YCSB_KEY VARCHAR(255) PRIMARY KEY,  
    FIELD0 TEXT, FIELD1 TEXT,  
    FIELD2 TEXT, FIELD3 TEXT,  
    FIELD4 TEXT, FIELD5 TEXT,  
    FIELD6 TEXT, FIELD7 TEXT,  
    FIELD8 TEXT, FIELD9 TEXT  
);
```

Rezultati primerjalne analize zmogljivosti





Izvedba analize stičnih operacij

Priprava podatkov

```
CREATE TABLE ext (  
  ycsb_key VARCHAR(255) PRIMARY KEY,  
  value int NOT NULL);  
  
INSERT INTO ext  
SELECT  
  ycsb_key,  
  LTRIM(RIGHT(ycsb_key,5),0)::int % 10 value  
FROM usertable  
ORDER BY ycsb_key  
LIMIT 100;
```

Poizvedba

```
\timing  
  
SELECT u.ycsb_key, u.field4  
FROM usertable u  
INNER JOIN ext e  
  ON e.ycsb_key = u.ycsb_key  
WHERE e.value = 4;
```

Optimizacija stičnih operacij

- eksperimentalna zastavica
- vrstni red stičnih operacij
- / → ~800ms

```
SET experimental_force_lookup_join = true;
```

```
\timing
```

```
SELECT u.ycsb_key, u.field4  
FROM ext e  
INNER JOIN usertable u  
    ON e.ycsb_key = u.ycsb_key  
WHERE e.value = 4;
```

Ugotovitve

- slabši rezultati zmogljivostne analize
- slaba podpora stičnim operacijam
- slaba podpora obstoječim orodjem
- + zelo enostavno upravljanje
- + zadovoljiva kompatibilnost s PostgreSQL
- + aktivna skupnost
- ? druge obremenitve (TPC-C, ...)
- ? več vozlišč (3+)
- ? naslednje verzije (2.0.1+)

Priporočila

- min 2 GB in 1 jedro
- CPU > RAM
- več vozlišč
- Google NTP storitev
- SSD
- hitro omrežje
- porazdeljene in replicirane podatke
- transakcijske obremenitve
- globalne aplikacije
- oblak

Hvala za pozornost!

YCSB vrste obremenitve

- izbira ključev po Zipfovi distribuciji (A, B, C, F)
- izbira vedno zadnjega ključa (D)

Vrsta obremenitve	SELECT	UPDATE	INSERT
A	50%	50%	
B	95%	5%	
C	100%		
D	95%		5%
F	50%	50%	

Izolacijski nivoji

- Serializable
- zastareli Snapshot

Izolacijski nivo	Dirty reads	Non-repeatable reads	Phantoms
Read Uncommitted	X	X	X
Read Committed		X	X
Repeatable Read			X
Serializable			