

Cockroach - Distributed SQL DB

NoSQL baze podataka



Univerzitet u Novom Sadu
Fakultet tehničkih nauka

CockroachDB

- ▶ Zvanična dokumentacija
- ▶ Kompatibilna sa PostgreSQL bazom
- ▶ Može se koristiti bilo koji postgres Go driver



CockroachDB

Uvod

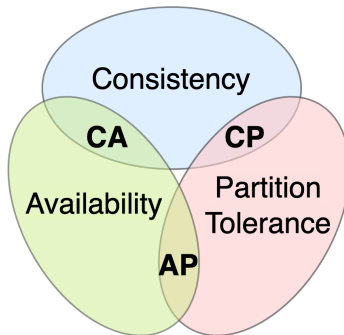
- ▶ Open-source
- ▶ Nastala iz ideje kreiranja open-source verzije *Spanner*-a
- ▶ Zašto takvo ime?
- ▶ Bubašvabe su izuzetno otporne
- ▶ Preživele katastrofu koja je dovela do izumiranja dinosaurus
- ▶ Kada prevedemo na bazu podataka...
- ▶ **Otpornost na otkaze**
- ▶ Kako postizemo otpornost na otkaze?

Osnovne informacije

- ▶ **Distribuirana SQL** baza podataka
- ▶ Podržava **SQL** upitni jezik
- ▶ Kreirana sa namerom da sadrži sve dobre osobine SQL baza podataka
- ▶ Ali i da usvoji dobre strane NoSQL baza podataka
- ▶ **Potpuna podrška za transakcije**
 - ▶ Nivo izolacije: **serijalizabilnost**
 - ▶ Potpuna podrška **ACID** operacija
- ▶ Podrazumevano **skalabilna = distribuirana**
- ▶ **CP = Consistency + Partitioning**

Distribuirana SQL baza podataka?

- ▶ Skalabilna SQL baza podataka
- ▶ Kako?



Distribuirana SQL baza podataka?

- ▶ Ne koristi relacioni pristup skladištu informacija
- ▶ Ne čuva podatke na disku kao tabele i redove
- ▶ **Koristi key-value skladište**

Kako funkcioniše?

Pogled nad bazom podataka na tri nivoa:

1. Nivo jezika

- ▶ API koji baza nudi
- ▶ SQL = Structured Query Language
- ▶ Poznat, prihvaćen, kompatibilan

2. Nivo izvršavanja upita

- ▶ **Magija**
- ▶ Kako baza izvršava upite nad podacima
- ▶ Mesto za vremenske i resursne **optimizacije**

3. Nivo skladišta

- ▶ **Key-value store**
- ▶ Posebna struktura ključeva
- ▶ Distribuirano skladište

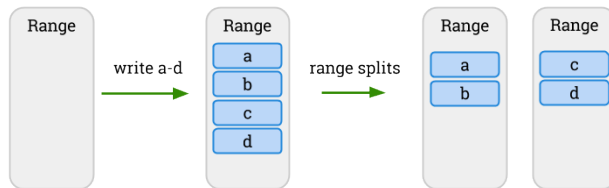
Nivo skladišta - struktura ključeva

- ▶ Mapiranje relacione tabele na key-value skladište
- ▶ ključ: $\langle \text{ime_tabele} \rangle / \langle \text{id} \rangle / \langle \text{ime_kolone} \rangle$
- ▶ vrednost: $\langle \text{vrednost_kolone} \rangle$

ID	<u>ISBN</u>	Title		ključ	vrednost
56	145-9-26-148410-0	The Hunger Games	→	knjiga/56/isbn	145-9-26-148410-0
				knjiga/56/naslov	The Hunger Games
7C	978-3-15-312410-1	The Great Gatsby	→	knjiga/7C/isbn	978-3-15-312410-1
				knjiga/7C/naslov	The Great Gatsby
9A	623-2-39-229320-2	1984	→	knjiga/9A/isbn	623-2-39-229320-2
				knjiga/9A/naslov	1984

Nivo skladišta - Range

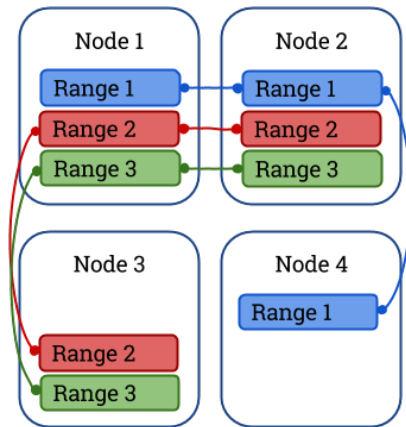
- ▶ Ključevi se ne generišu nasumično, već su deo *sortiranog skupa ključeva*
- ▶ Celokupno skladište se deli u **Ranges**
- ▶ Range je podskup sortiranog skupa ključeva, maksimalne veličine od 256 MB
- ▶ Kada se ceo podskup popuni, range se deli na dva manja
- ▶ Svaki Range je *uvek sortiran podskup početnog*
 - ▶ Pruža brze i jednostavne pretrage (indeksi slični *B-stablu*)
- ▶ Komunikacija između Range-ova se odvija putem *gRPC*-ja



Range Distribution & Rebalancing in CockroachDB

Nivo skladišta - Replikacija

- ▶ Range je jedinica replikacije
- ▶ Minimalni podrazumevani broj replikacija je 3
- ▶ Za održavanje konzistencije između replika se koristi **Raft algoritam**



Range Replication in CockroachDB

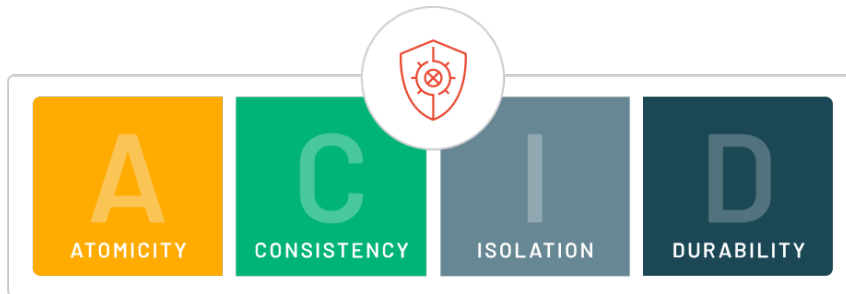
Preuzeto sa linka

Nivo izvršavanja upita - Raft

- ▶ Distribuirani koncenzus algoritam
- ▶ *Leader-follower princip*
- ▶ Leader se definiše za svaki Range
- ▶ Leader je zadužen za upise i propagaciju promena na replike
 - ▶ Glavna razlika u odnosu na *Gossip protokol* koji koristi Cassandra
 - ▶ U Gossip protokolu, svi čvorovi su jednaki
 - ▶ Bilo koji čvor može da obradi zahtev za upis i propagira informaciju svim replikama
- ▶ Predefinisano **CP baza podataka**
- ▶ Potrebno je da se uspostavi potvrda većine kako bi podatak bio validan
- ▶ Vizuelno objašnjenje rada Raft algoritma

Nivo izvršavanja upita - još magije

- ▶ Zvanična dokumentacija
- ▶ Detaljni opisi distribuiranih transakcija
- ▶ Sinhronizacije satova
- ▶ ACID
- ▶ ...



Nivo jezika

- ▶ SQL = Structured Query Language
- ▶ Kompatibilan sa PostgreSQL sintaksom
- ▶ *Dokumentacija*

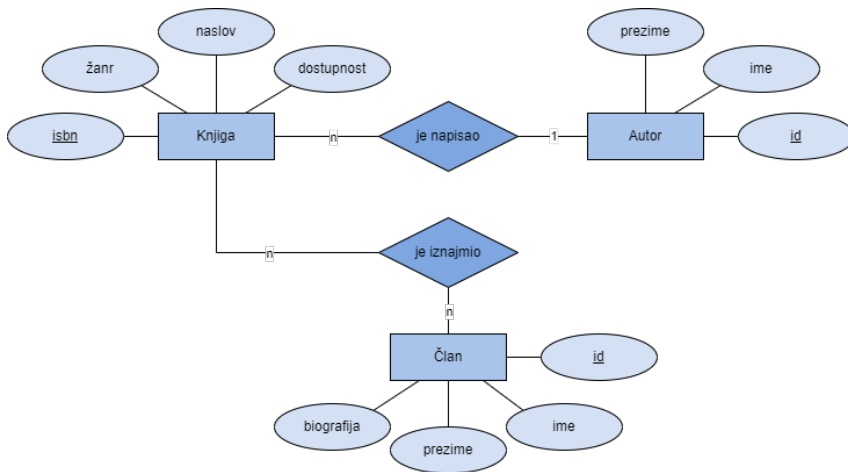


Primer

U okviru primera *RestCockroach* upotrebljeni su:

1. *Cockroach* - baza podataka
2. *Docker* - kontejnerizacija rešenja (i "instalacija" baza)
3. *Go* - implementacija primera
4. *Gorm* - objektno-relacioni mapper

Model podataka



Zadaci

- ▶ Proširiti servis tako da podržava:
 - ▶ dobavljanje svih knjiga jednog autora
 - ▶ dobavljanje svih knjiga koje je član iznajmljivao sa informacijama o tome da li je knjiga trenutno kod člana (zahteva dopunu modela)
 - ▶ vraćanje iznajmljene knjige
- ▶ **Bonus:** Istražiti razliku između Preload i Joins metoda