

5. Dokument baze podataka mongoDB

Thursday, 17 February 2022 19:00

Dokument baze podataka su baze cija namena je skladištenje, pribavljanje I upravljanje dokument-oriented ili polustrukturiranim podacima.

Polustrukturirani (nisu potpuni nestruktuirani) kako bi mogli da se obavljaju akcije nad njima (crud)

Key - value struktura gde je value dokument

Relacione I dokument baze (poredjenje)

Tabela	kolekcija dokumenata
Vrsta	dokument
Primarni kljuc	_id
Strani kljuc	DBRef
1:N relation	Aggregation
M:N relation	DBRefs array

Kod relacije se tabele spajaju svaki put I koci to

Kod dokument sve je u jednom dokumentu

Izgleda kod dokumenta nije definisano I fleksibilno je potpuno

Ako ne znate vrednost za nesto ne postoji to u bazi (ne zauzima mesto u memoriji)

Objekti se cuvaju kao dokument nema mismatch

Dokumenti su kompleksni koliko je to potrebno

Nezavisni su jedni od drugih

Otvoren format - podaci mogu da se cuvaju kao XML, JSON, BSON (binary JSON) itd.

Upisuje JSON I sam server konvertuje pre cuvanja u BSON jer je manji I efikasniji za pretragu

Tip koje podrzava JSON

Number, String, Boolean, Array, Object, null

Omogucava cuvanje verzije dokumenta - kao git

Brz pristup po kljucu za sve info u dokumentu

Visok nivo skalabilnosti

Podrska za replikacije

Organizacija dokumenta

Kolekcije - lista dokumenata istog tipa

Mehanizam tagova - dokumentima dodas tagove (kao da dokument pripada vise kolekcija)

Hijerarhije direktorijuma - kao file sistem

Nevidljivi metapodaci - kao datum kreiranja, korisnik koji je kreirao, ko je sta izmenio I kad ...

Eventual consistency - svi dokumenti nemaju isti update u svakom trenutku vec ima kasnjenje

Primena

Dinamicki podaci - kao wordpress

Polustrukturirani podaci

Web podaci - sesije, logovi, shopping cart
Obrada velike količine podataka

Problemi

Nema podrške za ACID transakcije

_id

Primarni ID se generise automatski

Sekundarni indeks je ono što pored primarnog definiše korisnik kao { ime: "Andrija"}

Ime je sekundarni

Db.kolekcija.find().explain() - explain objasni kako radi da bi optimisali sve

Razne komande za indeksiranje i traženje gledaj na prezentacijama

Memorija

Max memorija dokumenta 16mb

Za čuvanje velikih dokumenta metode

Files - metapodatke o dokumentu

Chunks - deli dokument na delove 256kb (kao lista delova koje čine dokument)

MongoDB nije dobar za čuvanje slika i videa

Replikacija

Jedna se kopija nazove master i kontrolise ostale slave replike

Sekundarne su samo čitanje

Master je upis pa se sinkuje sa slave replikama

Failover mehanizam - ako master propadne ostali glasaju novog mastera

Sharding mehanizam - deli se na delove dokument i deli na više replika

Jedna particija je opseg ključa

Kad particija pređe neku veličinu deli se na dve nove

Router prihvata zahtev od korisnika i rutira ga na instance od korisnika ka bazi

Map/Reduce

Map - cvor prihvata zahtev, postaje kordinator i deli task na male delove i daje ih worker cvorovima

Worker cvor može da uradi isto i da se napravi hijerarhija u obliku stabla

Worker obrađuje problem i šalje rezultat masteru

Reduce - skuplja rezultate worker cvorova i sastavlja ih u rezultat konačni