



UNIVERZITET U NOVOM SADU
FAKULTET TEHNIČKIH NAUKA



Sistemi baza podataka

MongoDB – UVOD

SADRŽAJ

- **NoSQL**
- Uvod u MongoDB
- Instalacija i pokretanje
- Studio 3T
- MongoDB dokumenti
- Uvod u rad sa dokumentima

NoSQL

- Not Only SQL – termin koji objedinjuje sve baze podataka i skladišta podataka koje ne slede primarne principe relacionih baza podataka.
- Motivacija za nastanak – eksplozija količine podataka koju je potrebno skladištiti – problem velikih količina podataka (*Big Data*)
 - potreba za visokim nivoom skalabilnosti,
 - potreba za visokim nivoom raspoloživosti.
- Dužina transakcije postaje problem – nepredvidiva dužina transakcije; **ACID** osobine transakcija više ne odgovaraju zahtevima vezanim za obradu velikih količina podataka.

Eksplzija količine podataka

2020 *This Is What Happens In An Internet Minute*

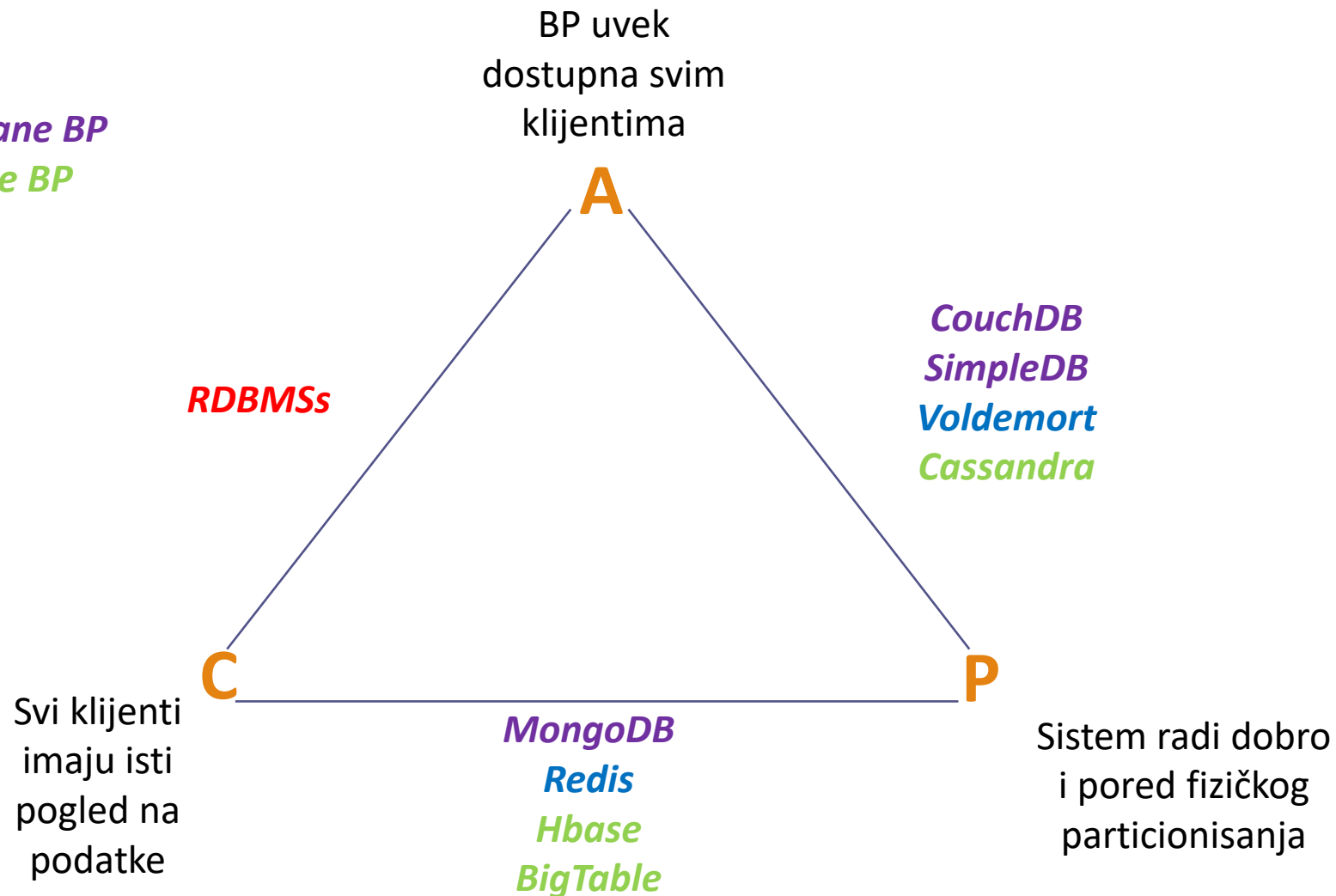


CAP teorema

- Sistem koji skladišti deljene podatke ne može obezbediti istovremeno zadovoljenje sledećih uslova:
 - konzistentnost (eng. *Consistency*) – svako čitanje iz baze podataka kao rezultat ima najnoviju verziju podataka (sve replike sadrže iste podatke);
 - raspoloživost (eng. *Availability*) – odziv sistema u garantovanim vremenskim okvirima, tj. baza podataka će uvek biti dostupna;
 - tolerancija razdvojenosti (eng. *Partition tolerance*) – nijedan skup otkaza (pri komunikaciji), osim potpunog otkazivanja, ne sme da proizvede neispravan odziv sistema baze podataka.
- **Primenjiva na sisteme zasnovane na distribuiranoj arhitekturi.**

CAP teorema

- *Relacione BP*
- *Ključ-vrednost BP*
- *Dokument-orijentisane BP*
- *Kolonski-orijentisane BP*



SADRŽAJ

- NoSQL
- **Uvod u MongoDB**
- Instalacija i pokretanje
- Studio 3T
- MongoDB dokumenti
- Uvod u rad sa dokumentima

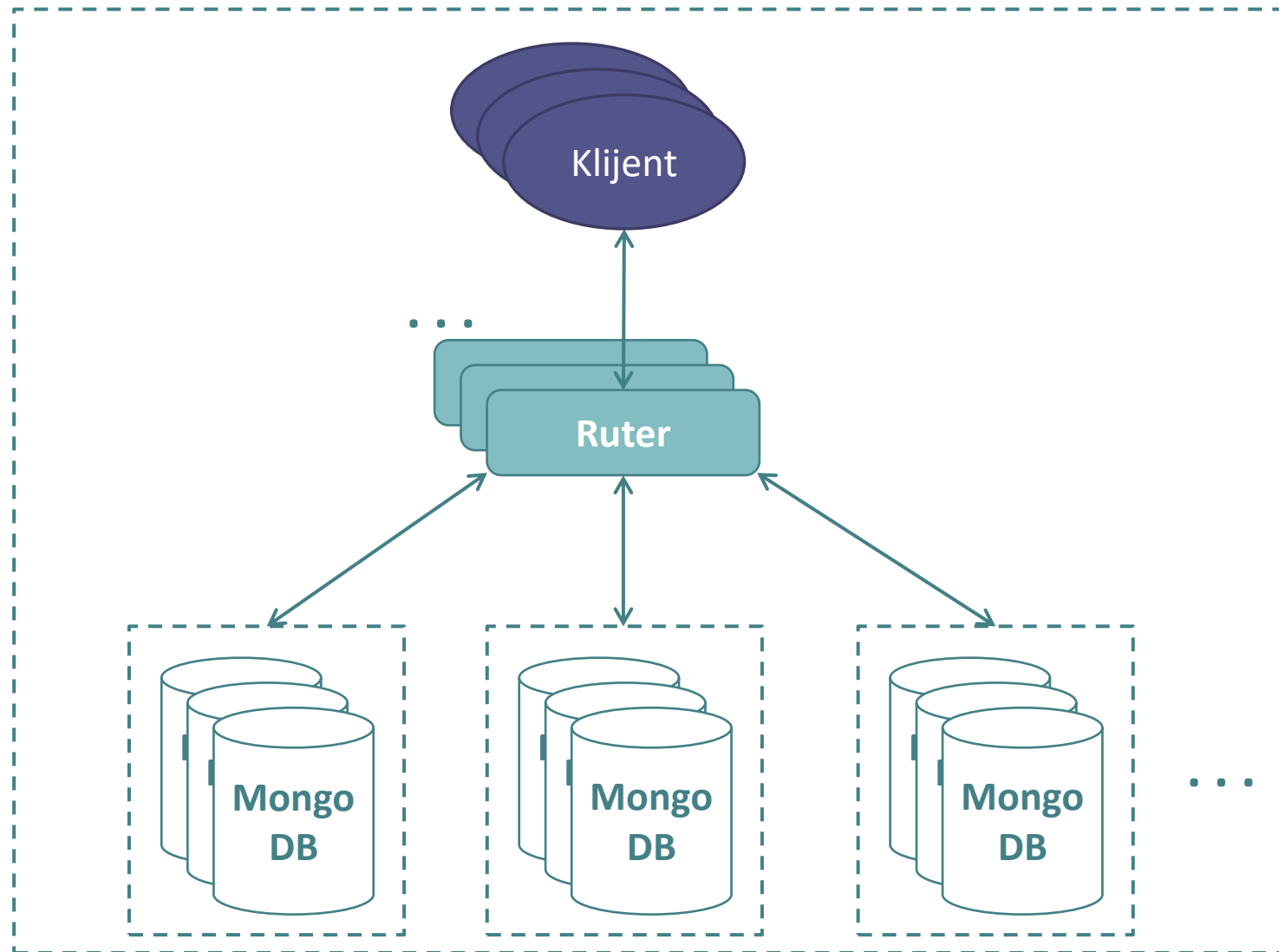
MongoDB

- MongoDB predstavlja open-source dokument-orijentisanu bazu podataka.
- Oblasti primene:
 - <https://www.mongodb.com/use-cases>
 - Pri radu sa nestrukturiranim i polustrukturiranim podacima
 - *Content management* (npr. *E-Commerce*), logovanje podataka i tome slično;
 - Pri ubrzanom razvoju prototipova (ne zahteva upotrebu migracija pri izmeni šeme bp).
- Zapise u MongoDB predstavljaju dokumenti čija struktura podataka je sačinjena od parova ključ-vrednost – MongoDB dokumenti su vrlo slični JSON objektima.

Zašto dokumenti?

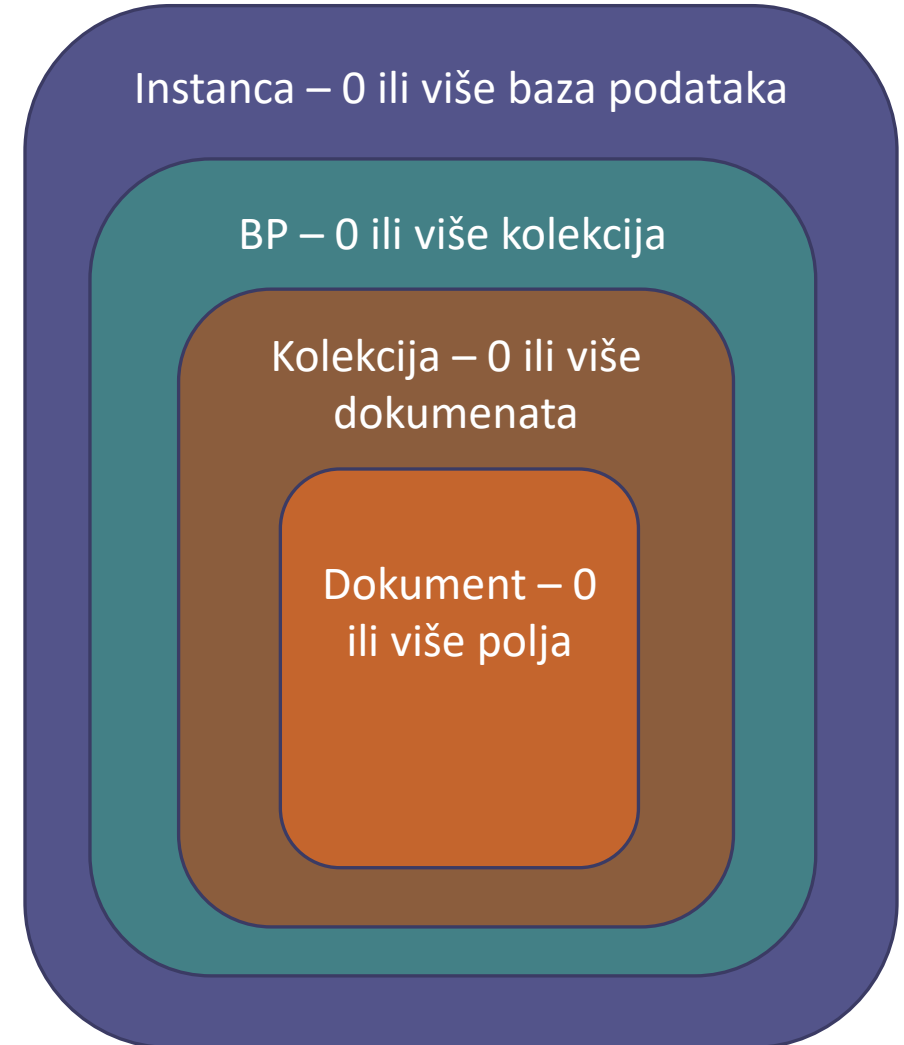
- Skladišta dokumenata su vrlo fleksibilna – dobro rukuju polustrukturiranim i nestrukturiranim podacima.
- Dobar su izbor kada nije unapred poznato kakva će biti struktura podataka koji će se skladištiti.
- Korisnici imaju mogućnost da kreiraju dokumente određene strukture bez uticaja na prethodno kreirane dokumente.
- Šema se može izmeniti bez prekida u radu servera, što vodi visokoj dostupnosti instance (*high availability*).

Primer MongoDB klastera



Hijerarhija objekata u MongoDB

- Jedna MongoDB instanca može imati 0 ili više baza podataka.
- Jedna baza podataka može imati 0 ili više kolekcija.
- Jedna kolekcija može imati 0 ili više dokumenata.
- Jedan dokument može imati 0 ili više polja.



Odnos prema konceptima relacione baze podataka

| RDBMS | | MongoDB |
|----------------|---|---------------------|
| Baza podataka | ➡ | Baza podataka |
| Tabela, pogled | ➡ | Kolekcija |
| Red, torka | ➡ | Dokument (BSON) |
| Kolona | ➡ | Polje |
| Indeks | ➡ | Indeks |
| Spoj | ➡ | Ugnježdeni dokument |
| Strani ključ | ➡ | Referenca |
| Particija | ➡ | <i>Shard</i> |

- Kolekcija ne nameće šta je to što bi trebalo da sadrži. U njih se skladište tematski (a ne strukturalno) slični dokumenti.
- Podaci nemaju eksplicitno iskazanu šemu (*schema-less*; šema je dinamička).
- Ugnježdeni dokumenti – ugnježdeni objekti.

Šema podataka u MongoDB

- MongoDB ne zahteva postojanje predefinisane šeme podataka
 - Ne postoji jezik za definisanje šeme podataka (*Data Definition Language*).
 - Svaki dokument u okviru jedne kolekcije može imati različit skup polja – šema je dinamička.

```
{  
  name: "Will"  
  birthplace: "NY",  
  aliases: [ "Bill",  
    "Mike" ],  
  loc: [ 32.7, 63.4 ],  
  boss: "Ben",  
  salary: 34,000.00  
}
```

```
{  
  name: "Ben"  
  birthplace: "AL",  
  boss: "Sue"  
}
```

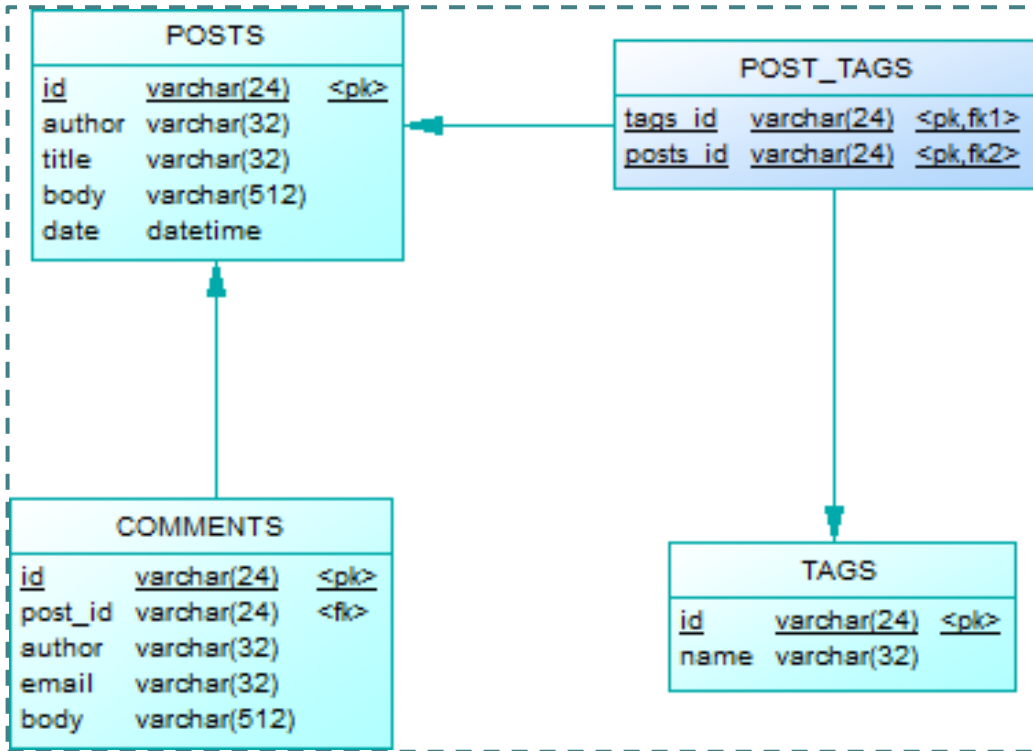
```
{  
  name: "Sue"  
}
```

```
{  
  name: "Jean"  
  birthplace: "WA",  
  aliases: [ "Billy" ],  
  loc: [ 38.5, 64.3 ],  
  boss: "Ben"  
}
```

Šema podataka u MongoDB

- Aplikacijom vođena šema (*Application Driven Schema*).
 - Naglasak se stavlja na šablone pristupa podacima (*data access patterns*), odnosno, na način na koji aplikacija koristi podatke.
 - Podaci se skladište tako da se podrži njihovo efikasno korišćenje – podaci koji se često koriste zajedno se ne razdvajaju, kako bi se izbegla potreba za spajanjem podataka.
 - MongoDB do verzije 3.2 nije ni podržavao mehanizam spajanja (*join*). Od verzije 3.2 uveden operator *\$lookup*, koji nameće različita ograničenja pri spajanju.
- Iako se šema u MongoDB ne deklariše eksplicitno, aplikacije u najvećem broju slučajeva koriste podatke na način kao da šema postoji.

Poređenje relacione šeme i MongoDB dokumenata



```
{
  "_id": ObjectId("cd4o5pf03dc53234efd23rs"),
  "author": "someone",
  "title": "Test title this is.",
  "body": "Test body this is.",
  "comments": [
    {
      "body": "comment body 1",
      "email": "johnDoe@doe.com",
      "author": "John Doe"
    },
    {
      "body": "comment body 2",
      "email": "janeDoe@doe.com",
      "author": "Jane Doe"
    }
  ],
  "date": ISODate("2018-01-07T08:58:58.941Z"),
  "tags": [ "mongoDB", "document", "difference" ]
}
```

SADRŽAJ

- NoSQL
- Uvod u MongoDB
- **Instalacija i pokretanje**
- Studio 3T
- MongoDB dokumenti
- Uvod u rad sa dokumentima

Instalacija

- Preuzeti odgovarajući [instalacioni fajl](#) sa zvanične internet stranice (MongoDB Community Sever).
 - Poslednja stabilna verzija – 4.4.5.
- Odabrati kompletnu instalaciju (unapred određeni parametri instalacije).
- Po želji isključiti/uključiti opciju za instalaciju MongoDB Compass klijentske aplikacije (na vežbama će biti korišćen Studio 3T).
- Nakon instalacije, izvršne datoteke za pokretanje procesa od interesa nalaze se u C:\Program Files\MongoDB\Server\4.4\bin direktorijumu.
 - Kako bi se izvršni fajlovi mogli pokretati kroz interfejs komandne linije (*Command Prompt*), potrebno je adresu direktorijuma u kom se nalaze izvršni fajlovi [dodati u Path promenljivu radnog okruženja \(Environment variable\)](#).
- Skinuti MongoDB [Database Tools](#) skup korisnih alata, te sadržaj skinutog bin foldera prekopirati u C:\Program Files\MongoDB\Server\4.4\bin.

Komponente MongoDB paketa

- Uz instalaciju MongoDB dolaze različite komponente – procesi:
 - *mongod* – instanca servera MongoDB.
 - Neophodno pokrenuti je pre rada sa podacima.
 - *mongo* – interaktivni klijent za rad sa podacima.
 - Ekvivalent **SQL*Plus** alatu kod *Oracle* baze podataka.
 - Predstavlja potpuno funkcionalno *JavaScript* okruženje.
 - *mongos* – proces koji je zadužen za realizaciju funkcionalnosti kod *shard*-ovanja.
 - *mongodump* – proces pomoću kog se vrši izvoz (*export*) sadržaja baze podataka.
 - *mongorestore* – proces pomoću kog se vrši uvoz (*import*) podataka generisanih od strane *mongodump* procesa.
 - Spisak svih komponenti može se pronaći u [zvaničnoj dokumentaciji](#).
- MongoDB procesi pokreću se pokretanjem njima odgovarajućih izvršnih fajlova (identičnog naziva).

Pokretanje instance

- Pokretanjem *mongod* izvršne datoteke pokreće se instanca MongoDB servera.
- Pri pokretanju je moguće specificirati vrednost različitih parametara instance, poput porta na kom sluša, korisničkog imena i lozinke, adrese radnog direktorijuma, *storage engine*-a itd.
 - Za spisak mogućih opcija uneti **mongod --help**.
- Kao podrazumevani port uzima se vrednost **27017**.
- Kao podrazumevana adresa radnog direktorijuma uzima se **C:\data\db**.
 - Ukoliko ovaj direktorijum ne postoji, neophodno ga je **ručno** kreirati (ukoliko se pokreće sa podrazumevanim parametrima).

SADRŽAJ

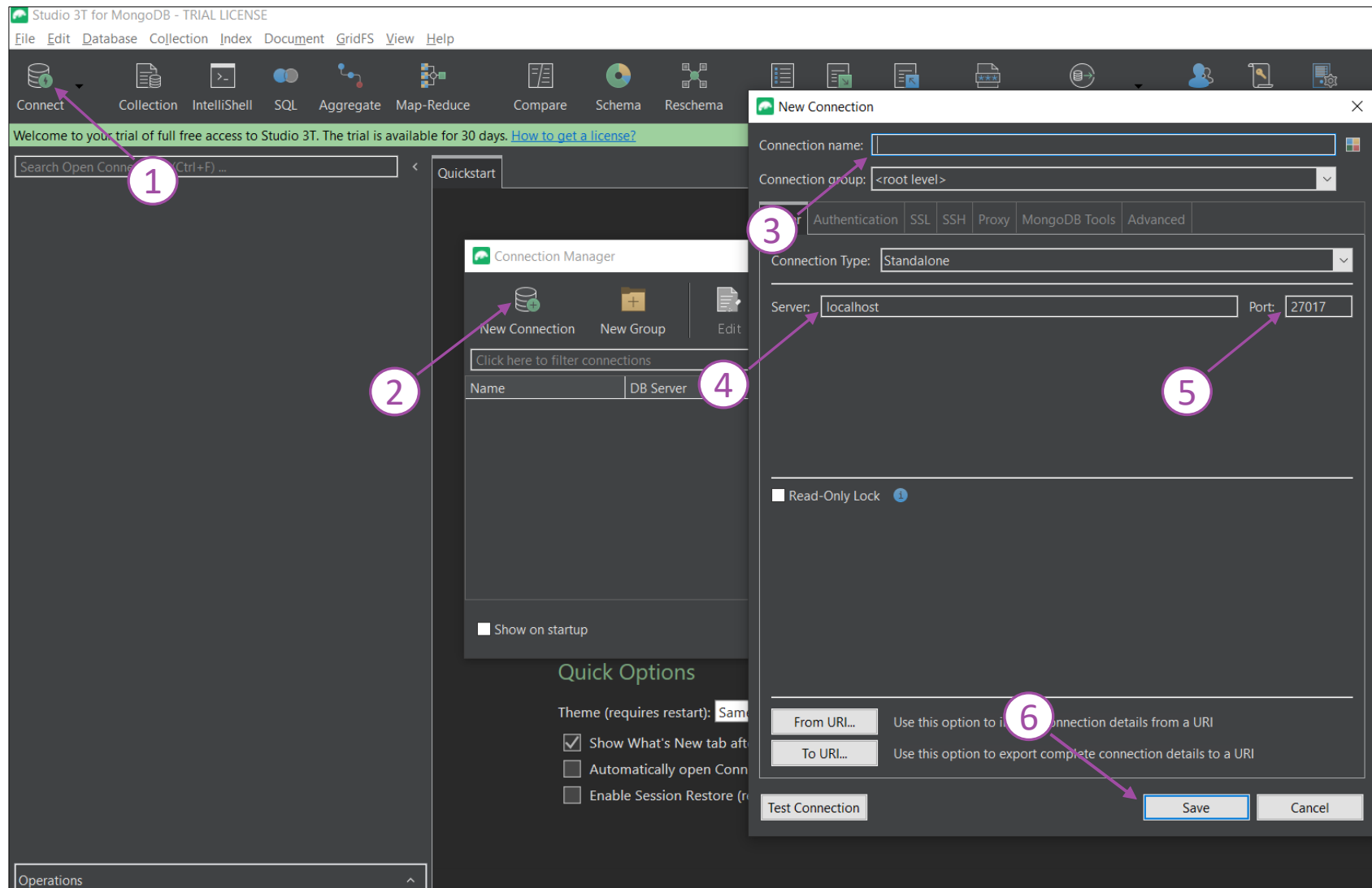
- NoSQL
- Uvod u MongoDB
- Instalacija i pokretanje
- **Studio 3T**
- MongoDB dokumenti
- Uvod u rad sa dokumentima

Studio 3T

- Klijentska aplikacija koja olakšava rad sa instancama servera MongoDB-a.
- Ranije poznat pod nazivom Robomongo.
- Poslednja stabilna verzija – 2021.2.0.
- Ekvivalent *Oracle SQL Developer* alatu za *Oracle* bazu podataka.
- Instalacija se može preuzeti sa [zvaničnog sajta](#).

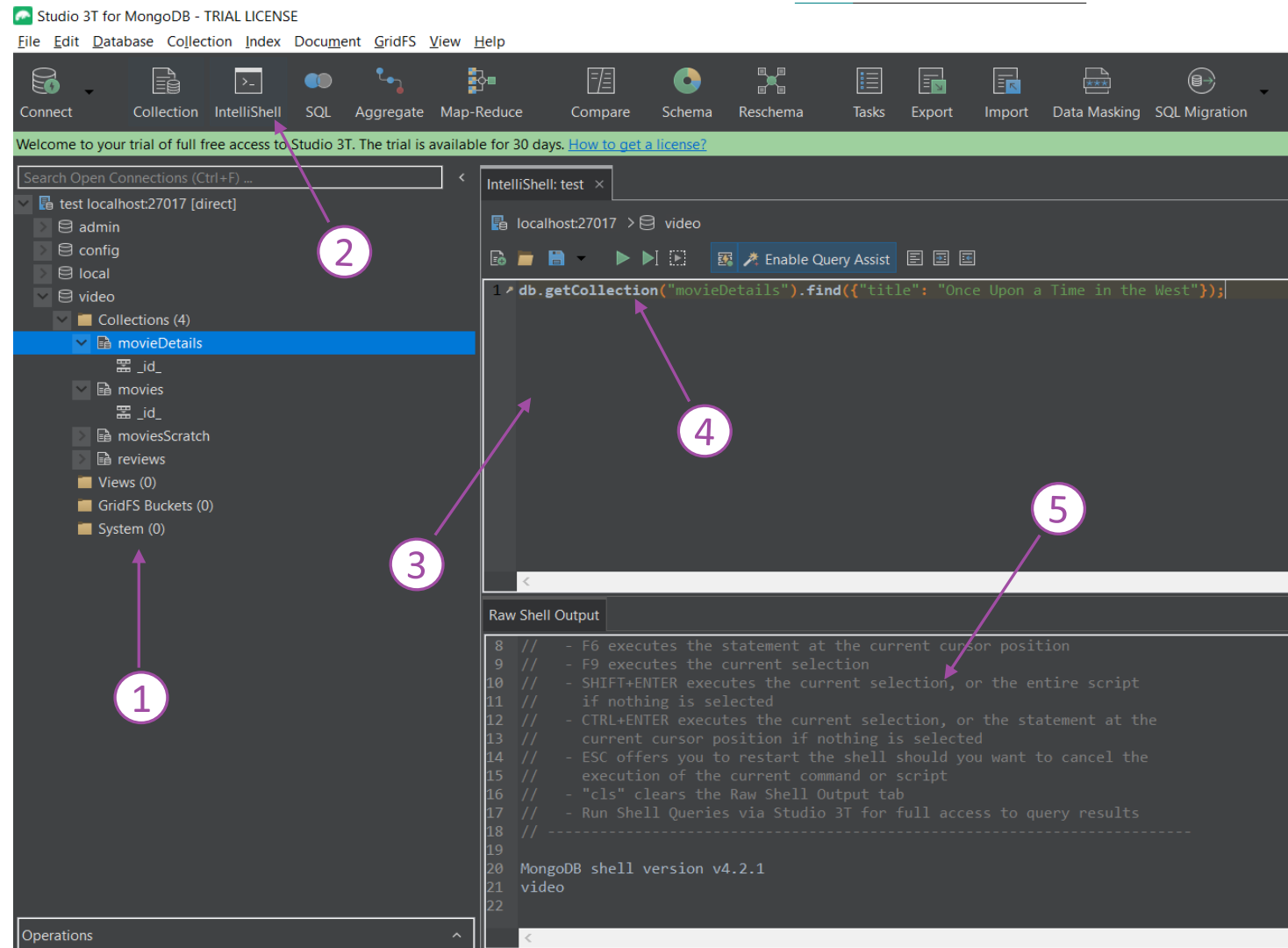
Studio 3T

1. Inicirati povezivanje,
2. odabrati kreiranje nove konekcije,
3. uneti naziv konekcije,
4. uneti adresu instance,
5. uneti port instance,
6. sačuvati konekciju.



Studio 3T

1. Prikaz trenutno postojećih baza podataka, sa njihovim unutrašnjim objektima;
2. Prozor za pokretanje naredbi (*shell*) se otvara odabirom IntelliShell opcije
3. Površina za unos naredbi
4. Pokretanje unetih naredbi (naredbe se mogu pokrenuti i sa *F5* ili *ctrl+enter*).
5. Površina za prikaz rezultata izvršavanja naredbe.



SADRŽAJ

- NoSQL
- Uvod u MongoDB
- Instalacija i pokretanje
- Studio 3T
- **MongoDB dokumenti**
- Uvod u rad sa dokumentima

JSON format podataka

- Polja JSON dokumenta/objekta smeštaju se unutar '{' i '}';
- Polja – parovi ključ-vrednost;
- Polja su međusobno razdvojena zarezom;
- Ključevi i vrednosti razdvojeni dvotačkom;
- Ključevi moraju biti stringovi;
- Vrednosti mogu biti tipa:
 - string,
 - broj,
 - Boolean,
 - objekat (ugnježdjeni) i
 - niz.
- Fleksibilnost koju JSON format podataka pruža omogućava podršku za različite šablona pristupa podacima
 - tako što omogućava kreiranje objekata koji sadrže sve podatke koji su potrebni kako bi se korisniku, na primer, prikazala *web* stranica.

```
{
  "first_name": "Paul",
  "surname": "Miller",
  "city": "London",
  "location": [45.123,47.232],
  "active": true,
  "address": {
    "street": "Kendal st",
    "number": 58,
    "zip_code": 56273
  },
  "cars": [
    {
      "model": "Bentley",
      "year": 1973,
      "value": 150000
    }
  ]
}
```

Koji su validni JSON dokumenti?

```
{  
  "name" : "Fred Flinstone";  
  "occupation": "Miner";  
  "wife": "Wilma"  
}
```

1

```
{ }
```

2

```
{  
  "title" : "Star Wars",  
  "quotes" : [  
    "Use the Force",  
    "These are not the droids  
you are looking for"  
  ],  
  "director" : "George Lucas"  
}
```

4

```
{  
  "city" = "New York",  
  "population" = 7999034,  
  "boroughs" = [  
    "queens",  
    "manhattan",  
    "staten island",  
    "the bronx",  
    "brooklyn"  
  ]  
}
```

5

```
{  
  "a" : 1,  
  "b" : {  
    "b" : 1,  
    "c" : "foo",  
    "d" : "bar",  
    "e" : [1, 2, 4]  
  }  
}
```

3

```
{  
  name : "Joe"  
}
```

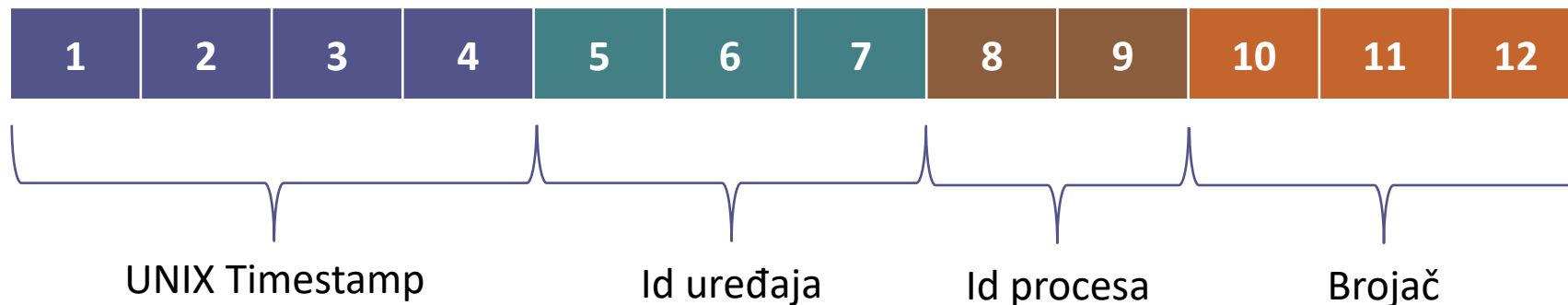
6

MongoDB dokumenti

- Zapise/torke u MongoDB predstavljaju dokumenti koji sadrže nula ili više polja.
 - Polja predstavljaju parovi ključ-vrednost.
- MongoDB dokumenti su vrlo slični dokumentima zapisanim u [JSON formatu](#).
 - Predstavljaju JSON dokumente u binarnom zapisu – BSON dokumenti.
 - Pri radu sad instancom preko Studio 3T korišćemo JSON.
 - MongoDB drajveri u višim programskim jezicima (JAVA, C#) dozvoljavaju neki vid rada sa JSON/BSON dokumentima (pored objektnog pristupa MongoDB objektima).
- Svaki MongoDB dokument **mora** posedovati jedinstvenu (na nivou kolekcije) vrednost „_id“ polja.
 - Ukoliko korisnik ne unese vrednost za ovo polje dokumenta, to će umesto njega učiniti drajver ili sama instanca (polje će biti automatski generisano).
 - Ukoliko se polje automatski generiše, vrednost polja se generiše korišćenjem objekta [ObjectId](#) klase.

Polje _id

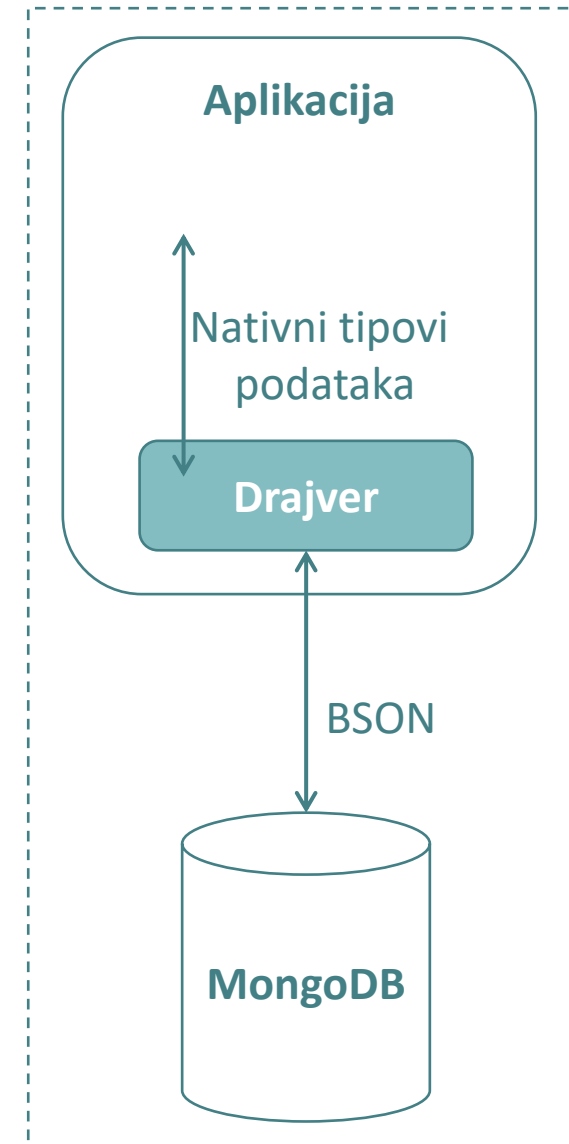
- Svaki MongoDB dokument **mora** posedovati jedinstvenu (na nivou kolekcije) vrednost „_id“ polja.
 - Ukoliko korisnik ne unese vrednost za ovo polje dokumenta, to će umesto njega učiniti drajver ili sama instanca (polje će biti automatski generisano).
 - Ukoliko se polje automatski generiše, vrednost polja se generiše korišćenjem objekta [ObjectId](#) klase.
 - Pri automatskom generisanju vrednosti ovog polja za dokument, generiše se dvanaestobitna vrednost koja ima sledeću strukturu:



Primer: "_id" : ObjectId("50906d7fa3c412bb040eb577")

BSON format podataka

- MongoDB zapravo skladišti podatke u [BSON formatu](#), odnosno, u vidu binarnog JSON-a.
- MongoDB drajveri šalju i prihvataju podatke u BSON formatu. Na aplikativnom nivou, oni mapiraju podatke zapisane u BSON formatu na odgovarajuće native tipove koje podržava programski jezik u kom je aplikacija napisana.
- BSON je dizajniran tako da:
 - zauzima manju količinu memorije od JSON formata,
 - je prolazak kroz elemente dokumenta olakšan,
 - prevođenje iz i u njegov format bude što je moguće efikasnije.
- BSON podržava veći broj tipova podataka u odnosu na JSON (datume; *int*, *float* itd. umesto samo *number*, binarne tipove za podršku skladištenju multimedijalnog sadržaja...).

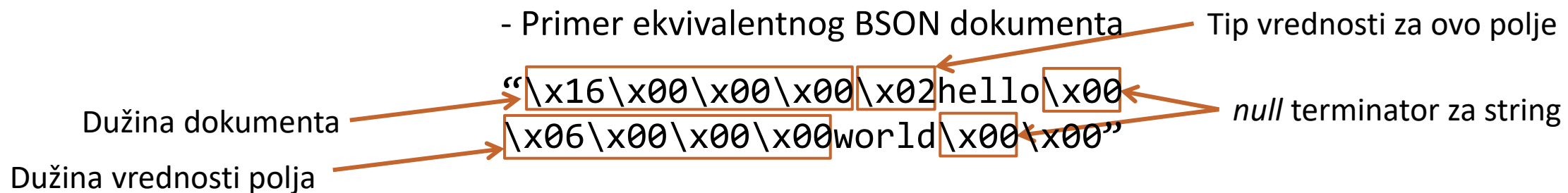


BSON format podataka

- Primer JSON dokumenta

```
{ "hello": "world" }
```

- Primer ekvivalentnog BSON dokumenta



SADRŽAJ

- NoSQL
- Uvod u MongoDB
- Instalacija i pokretanje
- Studio 3T
- MongoDB dokumenti
- **Uvod u rad sa dokumentima**

MongoDB naredbe

- U MongoDB postoje tri različite vrste naredbi nad bazom podataka:
 - Korisničke naredbe:
 - CRUD naredbe,
 - agregacione naredbe,
 - *geospatial* naredbe.
 - Operacije nad bazom podataka:
 - naredbe za autentikaciju, upravljanje korisnicima i ulogama (*role*), rukovođenje sesijama,
 - naredbe za replikaciju, *Sharding*,
 - naredbe za dijagnostiku i administraciju.
 - Naredbe za praćenje rada baze podataka (*Auditing*).
- Kompletan spisak naredbi može se pronaći [ovde](#).

Atomičnost operacija

- Sve operacije pisanja u MongoDB bazu podataka su **atomične na nivou jednog dokumenta**.
 - Kada jednom operacijom pisanja modifikujemo nekoliko dokumenata, modifikacija svakog dokumenta je atomična operacija, ali operacija pisanja u celini nije atomična – može doći do konflikata.
 - Od verzije 4.0 uvedena podrška i za *multi-document* ACID transakcije, ali njihovo korišćenje ima uticaj na performanse.

SADRŽAJ

- NoSQL
- Uvod u MongoDB
- Instalacija i pokretanje
- Studio 3T
- MongoDB dokumenti
- Uvod u rad sa dokumentima

PITANJA?

