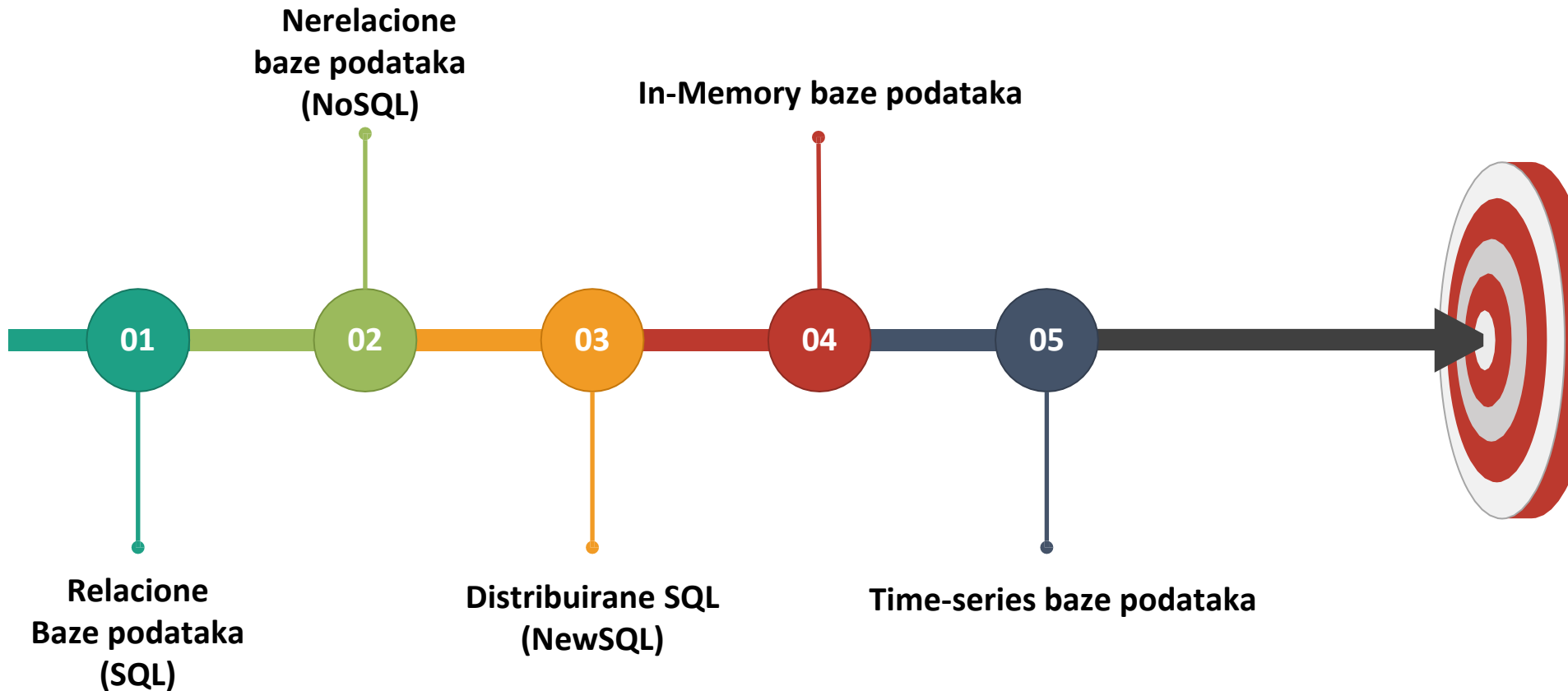


Tehnologije u bazama podataka

Predmet: Administriranje Baze Podataka
Predavač: dr Dušan Stefanović

Tehnologije u Bazama podataka



Tradicionalne relacione baze podataka SQL

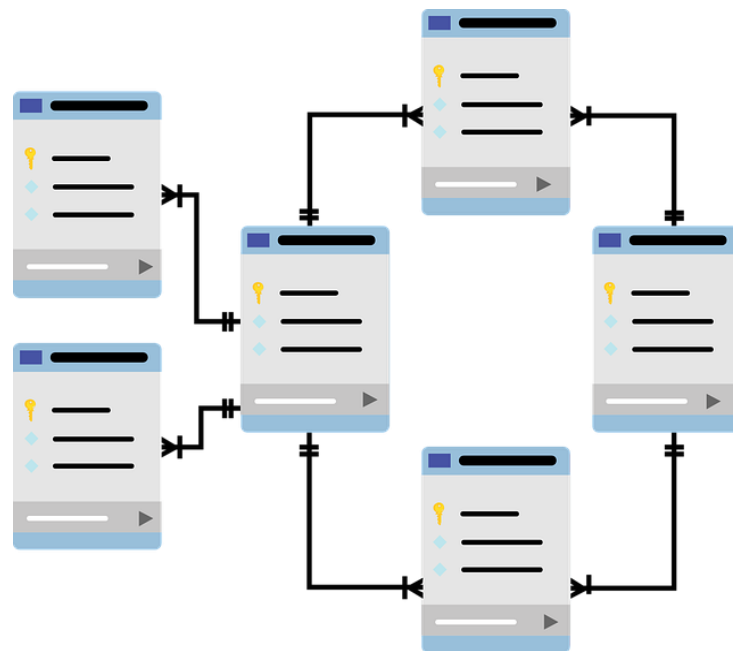
- **Relacione** baze podataka čuvaju podatke u tabelama

- Tabela se sastoji iz redova i kolona
- Svaka instanca baze podataka ima unapred **predefinisano šemu**
- Jedna ili više kolona u svakoj tabeli se koriste kao

Primarni ključ

- Redovi u jednoj tabeli mogu biti povezani sa redovima u drugoj tabeli preko

Stranog ključa



Baza podataka Online Shop

Klijent



Primarni ključ

Proizvod Id, Opis, Cena, Količina....

Proizvod



Porudžbine



Strani ključ

Strani ključ

Primarni ključ

Primarni ključ

Porudžbina Id, Datum, Klijent Id, Proizvod Id

Klijent ID, Ime, Tel, Adresa...

Jezik u relacionim bazama podataka - SQL

- SQL - Structured Query Language
- Relacione baze podataka su najviše primenjene u radu sa operacionim podacima (transakcije)
- **Predefinisana šema** u bazi podataka sprečava greške i obezbeđuje da stvari budu organizovane
- Ključevi obezbeđuju visok nivo fleksibilnosti programeru u efikasnom pristupu podacima
- Popularne SQL baze podataka:



Use case

- **Enterprise kompanije**
 - Aplikacije koje su instalirane on-premise ili privatnom cloud-u (privatni datacenter)
- **Internet tržište – Web aplikacije**
 - Opisuju na milione korisnika (povećanje saobraćaja i do 200% u danu)
 - 24h dostupnost
 - Problem sa skalabilnošću
 - Nisu najbolje rešenje za ovakve zahteve

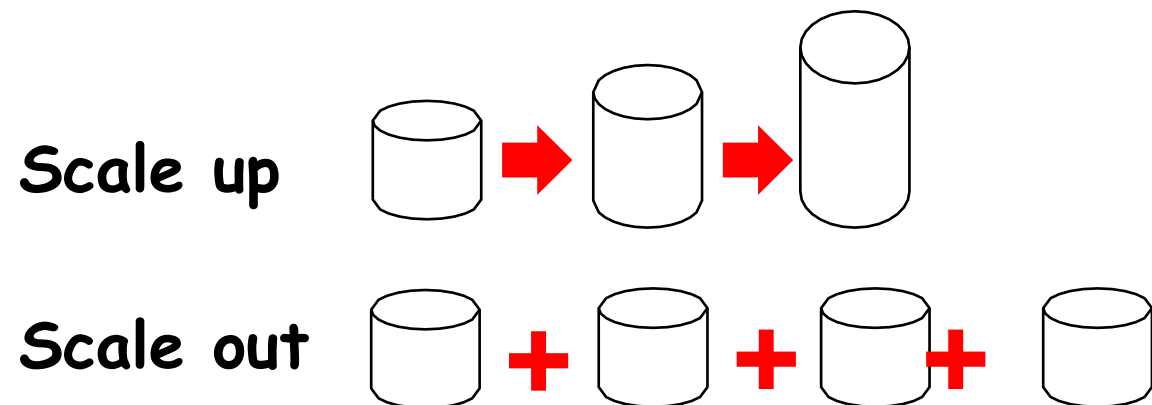


Najveći problem Relacionih baza podataka

- **Scalabilnost!**
 - Najveći problem kod relacionih baza podataka
 - Prilagođene su za **vertikalno skaliranje** (scale up/down)
 - Mašine koje su sve 'veće i veće' ali postoji ograničenje koliko jedan server može da se skalira
 - Prebacivanje na jači server je složen proces i obično zahteva period bez rada
- Web aplikacije zahtevaju dostupnost (*availability*) i *performanse*
 - Fokus je na alternativnim rešenjima NoSql bazama podataka...

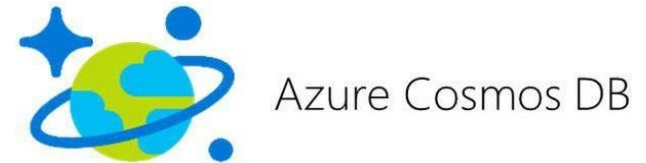
Nerelacione Baze podataka

- **Amazon, Facebook, Google....**
 - **Veliki** Internet igrači sa izazovima u radu sa **Velikim** podacima
 - *Scaling up* baze podataka je neefikasno i skupo
- **Ne-relacione Baze podataka**
 - Optimizovane su za web aplikacije
 - Fokus je na dve glavne stvari: **scale-out** i **Availability**



Nerelacione Baze podataka

- **Komercijalna i open-source NoSql baze podataka**



Nerelacione Baze podataka

- **Jednostavan Data Model**

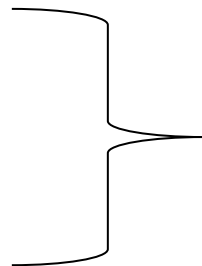
- Ne koriste relacioni model
- Nema strogo definisanu šemu
- Koristi jednostavan data model za čuvanje podataka u bazi koji je optimizovan za određene aplikacije (use case)

- **NoSQL Database** značenje

- "No support for SQL" – ne mogu da urade određene stvari koje SQL može
- "Not only SQL" a zatim je značenje promenjeno što znači da ne koriste relacioni model i strogo definisanu šemu

- **Vrste NoSQL baze podataka**

- *Key-value*
- Dokument
- *Wide column*
- Graf



Svaka je optimizovana za specifični use case

Nerelacione Baze podataka

PREDNOSTI

Šema
(Schema-on-Read)

SQL – Schema-on-Write
NoSQL – Schema-on-Read

Brzina
(Speed)

Fast Read/Write

Skaliranje
(Scaling Out)

SQL – Scaling Up
NoSQL – Scaling Up & Out

Dostupnost
(High Availability)

No single point of failure
Data replication

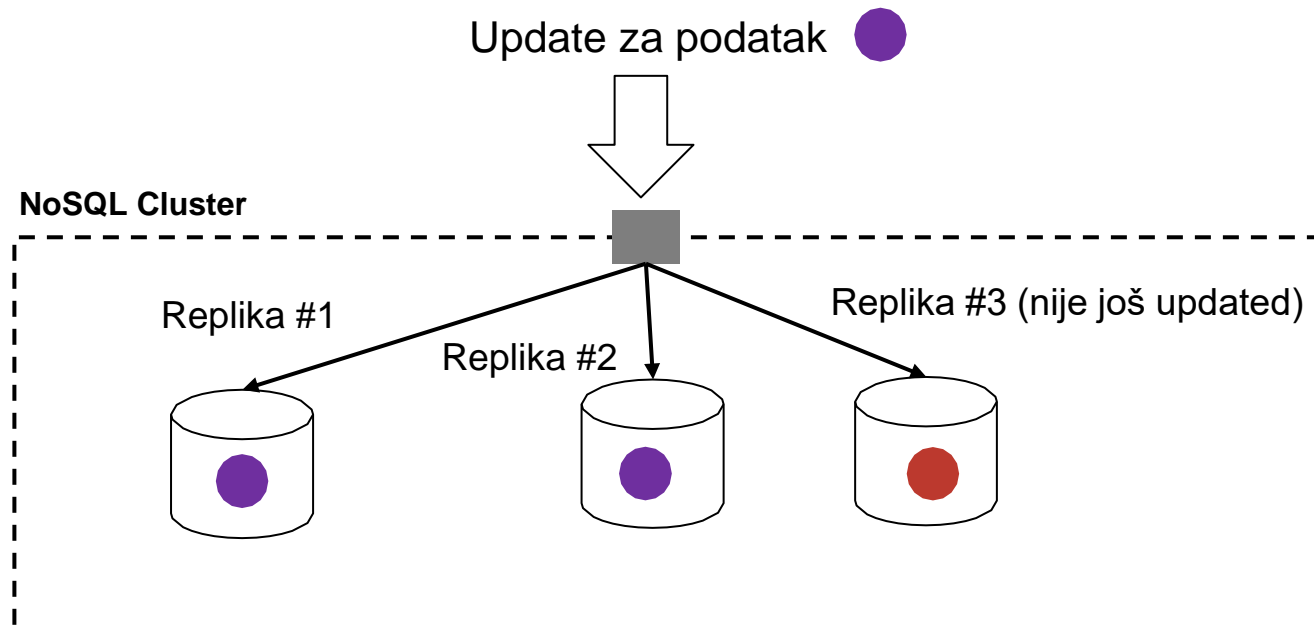
Nerelacione Baze podataka

NEDOSTACI

- **Nije transakciona !!!**
 - Ne ispunjava ACID model
- **Eventualni Konzistentni model (No Sql)**
 - “Slab” model konzistentnosti
 - Klaster baze podataka biće sinhronizovan sa istim podatkom nakon određenog vremena (**vremena konvergencije**)
 - Može da se desi da čitanje podataka ne bude konzistentno

▪ NoSQL Use Case

- Uglavnom su namenjene za analizu podataka



Tradicionalne Vs Nerelacione baze podataka

- You **win** something, you **lose** something...

- **Tradicionalne SQL baze podataka**

- **Ograničena skalabilnost i slaba otpornost** na otkaze
- Dobro se snalazi u radu sa **transakcijama** što je i najčešće zahtev u velikom broju aplikacija!

- **NoSQL baze podataka**

- Dizajnirane su za horizontalno skaliranje **scale out** koristeći distribuiranu arhitekturu
- Dostupnost (**Availability**) je bitnija od konzistentnosti
- Ograničena podrška za upravljanje transakcijama je veliki nedostatak
- Nije optimizovana za SQL upite

*We believe it is better to have application programmers deal with performance problems due to overuse of transactions as bottlenecks arise, rather than always **coding around the lack of transactions.***

(Google)

Distribuirane baze podataka (New SQL)

NewSQL baze podataka obezbeđuju da **SQL bude skalabilan**

NewSQL je klasa relacionog sistema za upravljanje bazama podataka da obezbedi skalabilnost Sql sistema za obradu velikog broja transakcija (OLTP) garantujući ACID model. (Wikipedia)

Relacioni
Data Model

Jaka
konzistentnost

SQL
funkcionalnosti

Horizontalna
skalabilnost

Visoka
dostupnost

Motivacija za razvoj distribuiranih baza podataka su nastale na osnovu dobrih osobina

Relacionih baza podataka

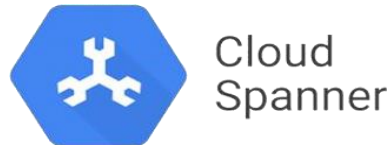
- relacioni model podataka
- čvrsta konzistentnost
- funkcionalnosti SQL-a

No Sql baza podataka

- horizontalna skalabilnost
- dostupnost u svakom trenutku

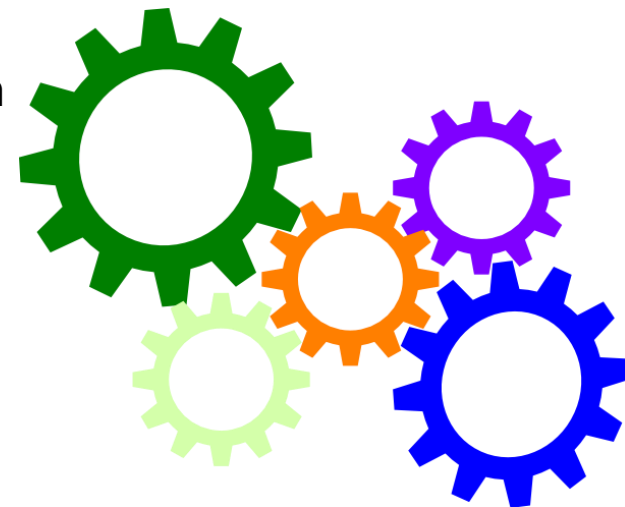
Distribuirane baze podataka (New SQL)

- Jedna **logička** relacionalna instanca koja je **distribuirana** i globalno skalirana na različite geografske regione
- Podaci su replicirani na više čvorova unutar klastera
- Horizontalno skaliranje klastera dodavanjem novih čvorova



Distribuirane baze podataka (New SQL)

- **NewSQL Ključne funkcionalnosti**
 - Podržava SQL (relacioni model)
 - Geografski distribuirana
 - Skaliranje po potrebi (vertikalno i horizontalno)
 - ACID saglasna (transakcije)
 - Načini instalacije (deployment)
 - Na lokaciji (on premises)
 - Self-hosted cloud (privatni cloud)
 - Baza kao servis (DBaaS)
 - Primena je na **Web scale aplikacije** sa analitikom u realnom vremenu i velikom frekvencijom transakcija



In Memory (IMDB) baze podataka

- Načini čuvanja podataka u bazi
 - **Disk-based baze podataka**
 - Tradicionalna i najčešće korišćena opcija
 - SSD/HDD tehnologije
 - **Podaci su trajni** (ne brišu se nakon restarta servera)
 - **Vreme pristupa disku** je usko grlo
 - **In-Memory baze podataka (IMDB)**
 - Podaci se čuvaju u memoriji (**in-memory**)
 - Mogu da budu SQL/NoSQL
 - **Prednost**
 - Uklanja se vreme potrebno da se traži podatak na disku
 - **Nedostatak**
 - Memorija (RAM) je znatno skuplja nego memorija na disku (disk storage) i samim tim na raspolaganju je manji prostor za čuvanje podataka

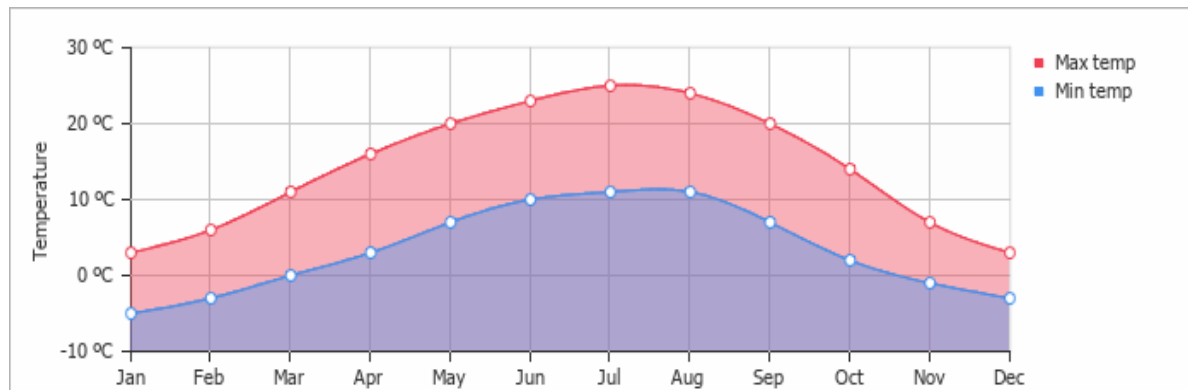
In Memory (IMDB) baze podataka

- Odlično rešenje za aplikacije koje zahtevaju **odgovor** iz baze na nivou **mikrosekunde**.
- Jedan od najčešćih slučaja primene in-memory baze podataka je keširanje podataka **data caching**



Time Series baze podataka (TSDB)

- Jedna od **brzo rastućih** kategorija baza podataka
 - Umreženi objekti koji non stop šalju podatke (data, data and data)
 - IoT uređaji i senzori, Monitoring IT...
 - Senzori koji očitavaju neke veličine...
 - Monitoring i Analiza
 - Veličina koja je bitna u takvim merenjima je **Vreme** (Timestamp)



Time Series baze podataka (TSDB)

Fokus je na Vreme!

- **Time-Series Data**
 - Niz ili serija vrednosti na **X osi koja predstavlja vreme**
 - Parametri koji se pamte
 - **Timestamp**
 - **Predmet npr. senzor**
 - **Izmerena vrednost**
 - Podaci mogu da se šalju periodično ili slučajno nakon nekog događaja
 - Radi se o podacima obima (high data volume)

Zahtevi od baze podataka

- Visoka učestalost upisa podataka
- Vremenske funkcije za programere
- **Primena tradicionalne relacione baze???**

Time Series baze podataka (TSDB)

- **Time Series Database (TSDB)**

- **Specijalizovana** baza podataka za podatke koji uključuju vreme u serijskim intervalima
- **Optimizovana** je dobro za skaliranje i upite reda ms.
- Automatski čuva trenutne podatke u in-memory za brzi pristup dok istorijske podatke čuva na disku
- Obraduje **high-volume data** (upis u bazu podataka)
- Ugrađene specijalizovane **time-related funkcije**

Optimizovane su za bolje rukovanje podacima koji sadrže vreme u serijskim intervalima (time series data)

