

Alternativni pristupi u izgradnji sistema baza podataka

NoSQL baze podataka

Sistemi baza podataka, dr Vladimir Dimitrieski

1

Sadržaj

- NoSQL baze podataka

2

2

NoSQL baze podataka

3

NoSQL baze podataka

- Termin *NoSQL*
 - nastao 2009. godine
 - u opštem smislu, objedinjuje sve baze podataka i skladišta podataka koje ne slede primarne principe relacionih baza podataka
 - akronim nema jasno definisano značenje
 - *No SQL* ili *Not Only SQL*
 - dobar *hashtag* na *Twitteru*
 - ne predstavlja jedan proizvod ili tehnologiju
 - već klasu proizvoda i kolekciju različitih, ponekad povezanih koncepata o načinima skladištenja i manipulacije podacima

4

4

NoSQL baze podataka

- Motivacija
 - pojava interneta prouzrokuje
 - eksploziju količine podataka koju je potrebno skladištiti
 - problem **velikih skupova podataka**
 - degradacija performansi RSUBP
 - količina podataka ima trend stalnog rasta
 - veličina pojedinačnih podataka takođe raste
 - potreba za visokim nivoom **skalabilnosti**
 - zahteva se unapređenje skalabilnosti
 - potreba za visokim nivoom **raspoloživosti**
 - eksploziju broja istovremenih zahteva nad BP
 - milioni istovremenih zahteva
 - odgovor mora biti u razumnom vremenskom roku

5

5

NoSQL baze podataka

- Motivacija
 - pojam **veliki skupovi podataka** predstavljaju kolekciju digitalnih informacija
 - engl. *big data*
 - čiji obim prevazilazi sposobnost prosečne percepcije ljudi pa i kapacitete raznih softverskih alata
 - za prihvatanje, obradu i upravljanje podacima
 - nije tačno određena donja granica veličine, definisane pojmom veliki podaci
 - trenutno - svaki skup podataka veći od nekoliko terabajta
 - istorijski prva i najčešće korišćena/citirana definicija koju je dao *Doug Laney* (često se naziva i 3 V definicija):
 - **veliki skup podataka je skup podataka koji sadrži raznovrsne podatke (*variety*) koji se prikupljaju u sve većem obimu (*volume*) i sve većom brzinom (*velocity*).**

6

6

NoSQL baze podataka

- Motivacija
 - veliki skup podataka je definisan kroz tri V osobine
 - **raznolikost podataka (engl. *variety*)** - dostupnost različitih tipova podataka u skupu
 - tradicionalni skupovi podataka su bili strukturirani sa jasno definisanim tipovima
 - danas, podaci se često prikupljaju u nestrukturiranom ili polustrukturiranom obliku
 - **količina podataka (engl. *volume*)** - količina prikupljenih podataka koju treba obraditi
 - najčešće je potrebno obraditi nestrukturirane podatke
 - u zavisnosti od kompanije, meri se u stotinama GB pa čak i u desetinama ZB

7

7

NoSQL baze podataka

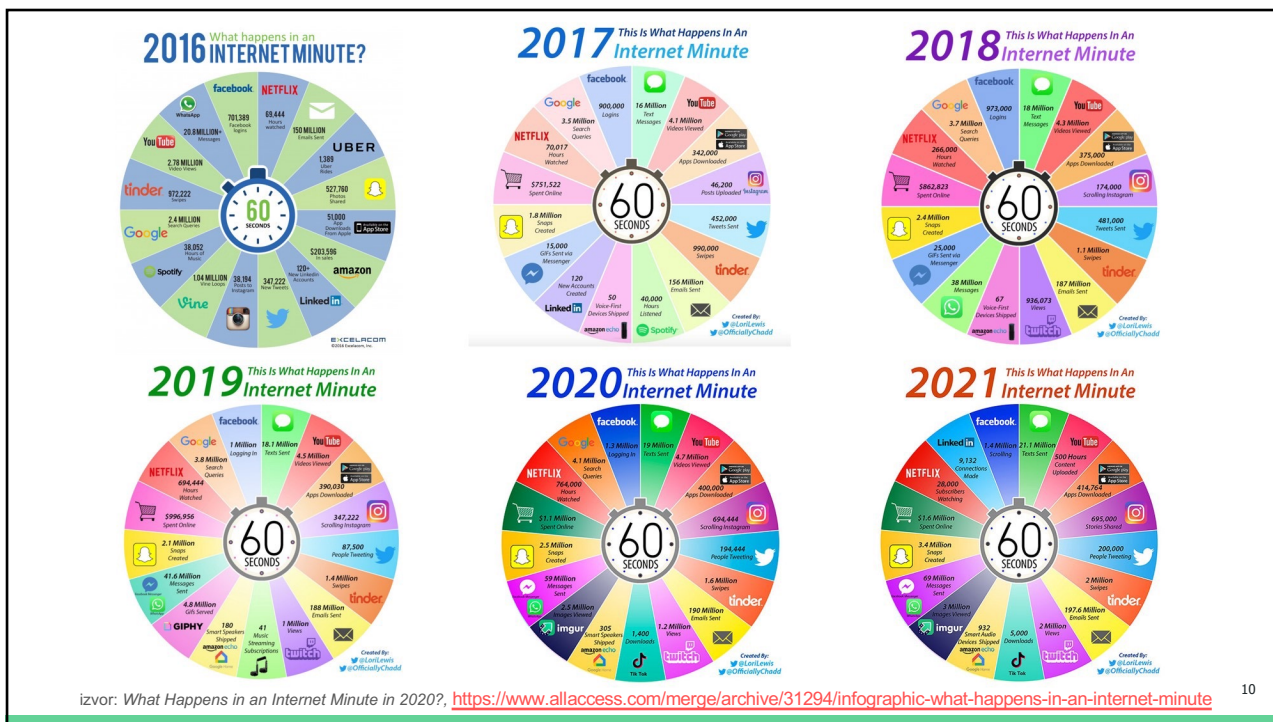
- Motivacija
 - veliki skup podataka je definisan kroz tri V osobine
 - **brzina prikupljanja podataka (engl. *velocity*)** - količina podataka koji su prikupljeni u jedinici vremena
 - nekada se posmatra i kao brzina po kojoj se prikupljaju i analiziraju podaci
 - najbrže prikupljanje podataka je ono koje smešta podatke u radnu memoriju

8

8

NoSQL baze podataka

- Motivacija
 - veliki skup podataka - dodatne osobine
 - **vrednost podataka (engl. value)** - vrednost koja može biti dobijena kao rezultat analize sprovedene nad prikupljenim podacima
 - karakteristika bitna za menadžment kompanije
 - vrednost koja im se može vratiti ulaganjem u prikupljanje i obradu velikih količina podataka
 - **verodostojnost podataka (engl. veracity)** - stepen istinitosti i tačnosti podataka
 - skup podataka može da sadrži i pogrešno/delimično prikupljene podatke, maliciozne podatke itd.
 - definiše da li su prikupljeni podaci odgovarajući za problem koji se rešava
 - od svih V osobina, **najteže je obezbediti verodostojnost podataka**



NoSQL baze podataka

- Motivacija
 - **skalabilnost** u tradicionalnim sistemima
 - mogućnost povećanja kapaciteta skladišta kroz dodavanje novih čvorova (računara) uz što manje gubitke performansi
 - jedan od osnovnih problema tradicionalnih sistema
 - u kojima se koriste relacione baze podataka
 - sa povećanjem količine podataka
 - više se brine o arhitekturi sistema
 - redovi čekanja, particije, replikacija itd.
 - nego o samoj poslovnoj logici
 - relacione baze podataka
 - nisu pogodne za horizontalno particionisanje
 - usled izuzetno loših performansi spojeva
 - mali broj relacionih baza podataka podržava interno horizontalno particionisanje

11

11

NoSQL baze podataka

- Motivacija
 - dužina transakcija postaje problem
 - nepredvidiva dužina transakcije
 - **ACID** osobine transakcija više ne odgovaraju zahtevima, vezanim za obradu velikih količina podataka
 - ACID osobine
 - skraćena od eng. *Atomicity, Consistency, Isolation, Durability*
 - skup osobina BP
 - garantuju pravilno izvršenje transakcija
 - definisane od strane Jim Gray-a
 - krajem 1970-ih godina
 - prilikom definisanja pouzdanog transakcionog sistema

12

12

NoSQL baze podataka

- ACID osobine
 - **atomičnost**
 - u slučaju uspešnog izvršenja transakcije
 - sve operacije obuhvaćene transakcijom će biti uspešno izvršene
 - baza podataka ostaje u konzistentnom stanju
 - u slučaju neuspešnog izvršenja transakcije
 - efekti izvršenja operacija obuhvaćenih transakcijom neće biti sačuvani u bazi podataka
 - baza podataka takođe ostaje u konzistentnom stanju
 - one, koje je prethodilo datoj transakciji

13

13

NoSQL baze podataka

- ACID osobine
 - **konzistentnost**
 - ukoliko su se podaci pre izvršenja transakcije nalazili u konzistentnom stanju, nalaziće se i po okončanju transakcije
 - **izolacija**
 - operacije tekuće transakcije ne utiču na druge transakcije
 - koje se istovremeno izvršavaju
 - **trajnost**
 - promene koje nastaju nakon završetka transakcije ostaju trajno sačuvane
 - čak i u slučaju pada čitavog sistema

14

14

NoSQL baze podataka

- CAP teorema
 - sistem koji **skladišti deljene podatke** ne može obezbediti istovremeno zadovoljenje sledećih uslova
 - **konzistentnost**
 - *eng. Consistency*
 - **raspoloživost**
 - *eng. Availability*
 - **tolerancija razdvojenosti**
 - *eng. Partition tolerance*
 - **primenjiva na sisteme zasnovane na distribuiranoj arhitekturi**

15

15

NoSQL baze podataka

- CAP teorema
 - **konzistentnost**
 - svako čitanje iz baze podataka kao rezultat ima najnoviju verziju podataka
 - **raspoloživost**
 - odziv sistema u garantovanim vremenskim okvirima
 - baza podataka će uvek biti dostupna
 - nezavisno kada je postavljen upit
 - postiže se
 - velikim brojem fizičkih servera
 - replikacijom podataka

16

16

NoSQL baze podataka

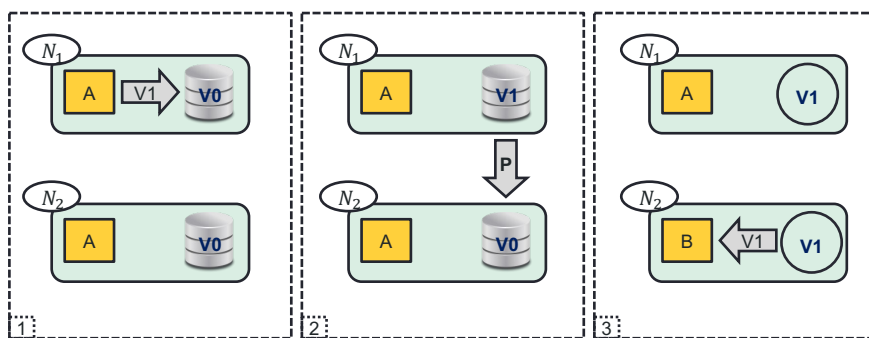
- CAP teorema
 - **tolerancija razdvojenosti**
 - nijedan skup otkaza, osim potpunog otkazivanja, ne sme da proizvede neispravan odziv sistema baze podataka
 - sistem mora da prihvata delimične otkaze i nastavlja ispravan rad
 - **razdvojenost**
 - stanje komunikacione mreže kod koje su delovi sistema podeljeni u dve ili više particija
 - između kojih ne postoji komunikacija
 - rešenje:
 - replikacija
 - odloženo (asinhrono) pisanje u trenutno nedostupne delove baze podataka

17

17

NoSQL baze podataka

- CAP teorema – dokaz



18

18

NoSQL baze podataka

- CAP teorema – dokaz
 - **šta se događa ako poruka P ne stigne na odredište?**
 - dolazi do razdvojenosti delova sistema
 - tri moguća događaja
 - **transakcija A se poništava**
 - sistem **ne prihvata razdvojenost** svojih delova
 - **transakcija A se uspešno nastavlja**
 - sistem **nije konzistentan**
 - **transakcija A čeka na uspešno slanje poruke P**
 - sistem **nije raspoloživ**
 - nepoznato vreme odaziva
 - u svakom od događaja **jedan od uslova nije ispunjen**

19

19

NoSQL baze podataka

- CAP teorema – kompromisi
 - **odbacivanje prihvatanja razdvojenosti**
 - obezbeđuju se konzistentnost i raspoloživost
 - alternative
 - kompletan sistem na jednom računaru
 - rad bez razdvojenosti
 - višestruko umrežavanje
 - redundantne veze
 - veoma skupe
 - ređe se pravi ovakav kompromis

20

20

NoSQL baze podataka

- CAP teorema – kompromisi
 - **odbacivanje raspoloživosti**
 - obezbeđuju se konzistentnost i tolerancija razdvojenosti
 - u slučaju razdvojenosti ne garantuje se vreme odziva
 - problem se redukuje uspostavljanjem što niže sprege između čvorova
 - koristi se veoma retko
 - sistem koji nije raspoloživ nije ni upotrebljiv

21

21

NoSQL baze podataka

- CAP teorema – kompromisi
 - **odbacivanje konzistentnosti**
 - obezbeđuju se raspoloživost i tolerancija razdvojenosti
 - **ne garantuje čitanje poslednje verzije podataka u slučaju razdvojenosti**
 - u suprotnosti sa ACID skupom osobina
 - definiše se **BASE** skup osobina
 - eng. *Basically Available, Soft state, Eventually consistent*

22

22

NoSQL baze podataka

- BASE osobine
 - **suštinski raspoloživ**
 - većina podataka je dostupna veći deo vremena
 - **nekonzistentno stanje**
 - baza podataka ne mora biti konzistentna u svakom trenutku
 - **konvergentna konzistencija**
 - dodatkom novog čvora u sistem sadržaj se može replicirati na njega
 - ne postoji garancija da će u svakom trenutku svi čvorovi sadržati identične kopije podataka
 - teži se vremenskoj tački u kojoj će svi čvorovi sadržati konzistentne podatke

23

23

NoSQL baze podataka

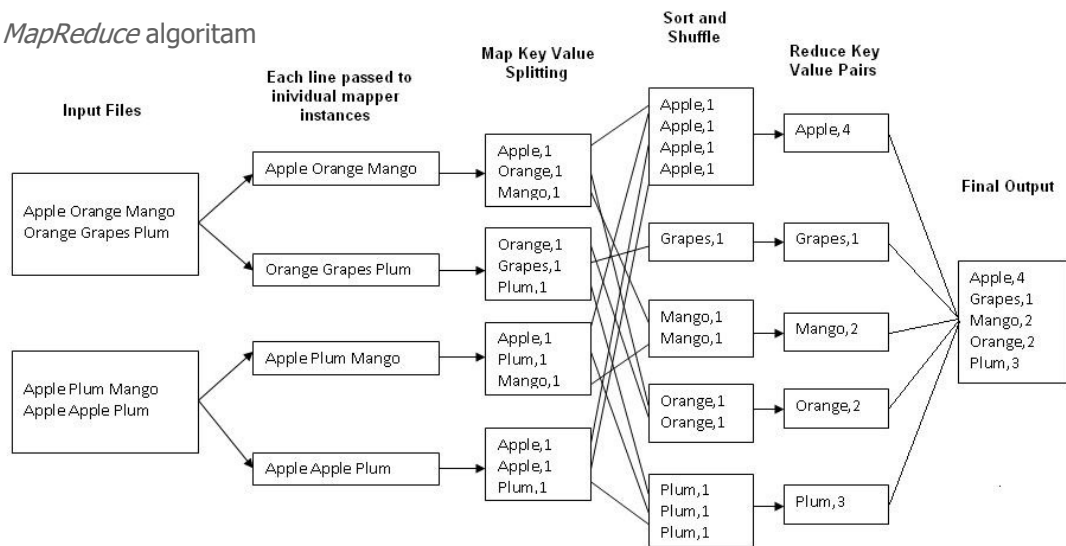
- *MapReduce* algoritam
 - omogućava upravljanje velikim količinama podataka
 - posredstvom distribuirane mreže čvorova
 - deli veliki zadatak na manje delove
 - pogodno za paralelno procesiranje
 - *MapReduce* okruženje
 - upravlja izvršenjem MR algoritma
 - obezbeđuje spregu sa korisnikom

24

24

NoSQL baze podataka

- *MapReduce* algoritam



26

26

NoSQL baze podataka

- NoSQL sistemi
 - nastali usled potreba za
 - skladištenjem velike količine podataka
 - pristupom od strane velikog broja korisnika
 - realizacija baza podataka se oslanja na distribuiranu arhitekturu
 - CAP teorema
 - sa BASE skupom osobina
 - uvodi pojam konvergentne konzistentnosti kao kompromisno rešenje
 - *MapReduce* algoritam
 - efikasna obrada podataka

27

27

NoSQL baze podataka

- NoSQL sistemi – osobine i prednosti
 - u opštem slučaju, **ne poseduju formalno specificiranu šemu baze podataka**
 - rukovanje fleksibilnim strukturama podataka
 - nestrukturirani i polustrukturirani podaci
 - **ne oslanjaju se na relacioni model podataka**
 - ne podržavaju operacije spajanja
 - **unapređene performanse u radu sa velikom količinom podataka**
 - **horizontalna skalabilnost i elastičnost**
 - povećanje kapaciteta i performansi u radnom režimu sistema
 - dodavanjem čvorova u distribuiranu mrežu

28

28

NoSQL baze podataka

- NoSQL sistemi – osobine i prednosti
 - **arhitektura lokalizovanih resursa**
 - engl. *shared-nothing architecture*
 - svaki server se oslanja na lokalnu masovnu memoriju
 - **particionisanje baze podataka**
 - ukupan broj zapisa je distribuiran u particije
 - dodeljene različitim serverima
 - particije se repliciraju
 - **asinhrona replikacija**
 - podaci nisu replicirani onog momenta kada su zapisani

29

29

NoSQL baze podataka

- NoSQL sistemi – osobine i prednosti
 - **izbegavanje objektno-relacionog mapiranja**
 - podaci se čuvaju u posebnim strukturama podataka
 - obično jednostavnim
 - bliskim objektno orijentisanim programskim jezicima
 - često predstavlja **sistem za upravljanje datotekama**
 - **visoka tolerancija na otkaze**
 - **daje prednost performansama**
 - na štetu konzistentnosti
 - **pristup putem API-ja** ili nekog drugog interfejsa
 - ne oslanja se na SQL
 - **dominantno otvorenog kôda**

30

30

NoSQL baze podataka

- NoSQL sistemi – nedostaci
 - **nepostojanje standarda**
 - **nepostojanje standardnog upitnog jezika**
 - UnQL (engl. *Unstructured Query Language*)
 - predlog upitnog jezika
 - zasnovan na JSON-u
 - **dominantno otvorenog kôda**
 - nedostatak dokumentacije i korisničke podrške

31

31

NoSQL baze podataka

- Podela NoSQL sistema
 - skladišta podataka tipa ključ-vrednost
 - kolonski orijentisano skladište podataka
 - baze podataka orijentisane ka dokumentima
 - baze podataka orijentisane ka grafovima

32

32

NoSQL baze podataka

- Skladišta podataka tipa ključ-vrednost
 - engl. *key – value stores*
 - svaki zapis se sastoji od skupa parova ključ – vrednost
 - proizvoljan broj parova
 - svojstva sa nedostajućim vrednostima se izostavljaju
 - **ključ** predstavlja naziv svojstva
 - jedinstveni identifikator koji ukazuje na vrednost
 - definiše se prilikom formiranja zapisa
 - **vrednost** sadrži podatak ili pokazivač na lokaciju koja sadrži podatak
 - u opštem slučaju vrednosti predstavljaju nizove karaktera
 - zapisi se čuvaju u okviru tabela

33

33

NoSQL baze podataka

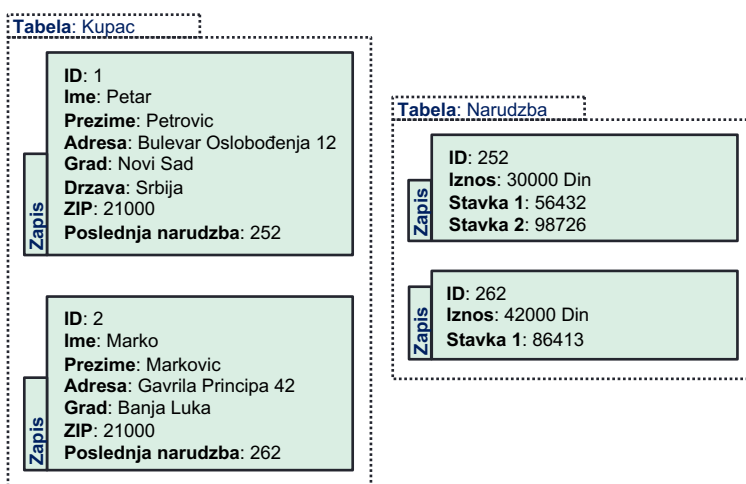
- Skladišta podataka tipa ključ-vrednost
 - zasnovane na principima struktura podataka u višim programskim jezicima
 - kolekcije
 - rečnici
 - rasute tabele
 - mogu da služe kao mehanizmi za keširanje

34

34

NoSQL baze podataka

- Skladišta podataka tipa ključ-vrednost



35

35

NoSQL baze podataka

- Skladišta podataka tipa ključ-vrednost

Ključ-vrednost	Riak	Tokyo Cabinet/Tyrant	Voldemort
Napisana u	Erlang	C	Java
Platforme	Linux	Linux	Bilo koja Java
Protokol	HTTP/REST	HTTP/memcached	HTTP/Thrift
Način čuvanja podataka	plug-in BP	RAM ili disk	RAM ili BerkeleyDB
Replikacija	asinhrona	asinhrona	asinhrona
Particionisanje	da	ne	da
Indeksi	da	ne	ne
MapReduce	da	ne	ne
Transakcije	ne	lokalne	ne

36

36

NoSQL baze podataka

- Kolonski orijentisano skladište podataka
 - engl. *wide column stores*
 - ne zahteva kompletnu šemu baze podataka
 - samo unapred definisane familije kolona
 - skup kolona koji predstavlja logičku celinu
 - kolone unutar familije ne moraju biti unapred definisane
 - visoke performanse prilikom vršenja upisa novih vrednosti

37

37

NoSQL baze podataka

- Kolonski orijentisano skladište podataka
 - trodimenzionalna struktura
 - identifikator reda
 - familija kolona i identifikator kolone
 - vremenska oznaka
 - revizija podataka
 - red u skladištu je analogan dokumentu
 - grupisani u kolekcije
 - primarni ključ zapisa služi za pristup redu
 - familije kolona igraju ulogu sekundarnog ključa

38

38

NoSQL baze podataka

- Kolonski orijentisano skladište podataka

Tabela: Kupac

ID: 1
CF/SC: Ime
 C: Ime: Petar
 C: Prezime: Petrovic
CF/SC: Adresa
 C: Adresa: Bul. Osl. 12
 C: Grad: Novi Sad
 C: Drzava: Srbija
 C: ZIP: 21000
CF/SC: Narudzbe
 C: Poslednja narudzba: 252

ID: 2
CF/SC: Ime
 C: Ime: Marko
 C: Prezime: Markovic
CF/SC: Adresa
 C: Adresa: Gavriła Principa 42
 C: Grad: Banja Luka
 C: ZIP: 21000
CF/SC: Narudzbe
 C: Poslednja narudzba: 262

Tabela: Narudzba

ID: 252
CF/SC: Cena
 C: Iznos: 30000 Din
CF/SC: Stavke
 C: Stavka 1: 56432
 C: Stavka 2: 98726

ID: 262
CF/SC: Cena
 C: Iznos: 42000 Din
CF/SC: Stavke
 C: Stavka 1: 86413

39

39

NoSQL baze podataka

- Kolonski orijentisano skladište podataka

Ključ-vrednost	Cassandra	Hbase	SimpleDB
Napisana u	Java	Java	Erlang
Platforme	Linux, Windows	Cross - platform	EC2, POSIX
Protokol	TCP/Thrift	HTTP/REST ili TCP/Thrift	HTTP/REST
Način čuvanja podataka	disk	Hadoop File System	S3
Replikacija	asinhrona	asinhrona	asinhrona
Particionisanje	da	da	da
Indeksi	da	ne	da
MapReduce	da	da	ne
Transakcije	lokalne	lokalne	ne

40

40

NoSQL baze podataka

- Baze podataka orijentisane ka dokumentima
 - engl. *document-oriented databases*
 - dokument
 - ekvivalentan redu u relacionim bazama podataka
 - u bazi podataka postoji niz samostalnih dokumenata
 - rezultat upita su podaci sadržani u jednom dokumentu
 - ne u povezanim tabelama
 - ne postoji potreba za šemom baze podataka
 - dokumenti u opštem slučaju ne moraju da imaju istu strukturu
 - dokumenti mogu da budu **međusobno povezani** putem URI-ja

41

41

NoSQL baze podataka

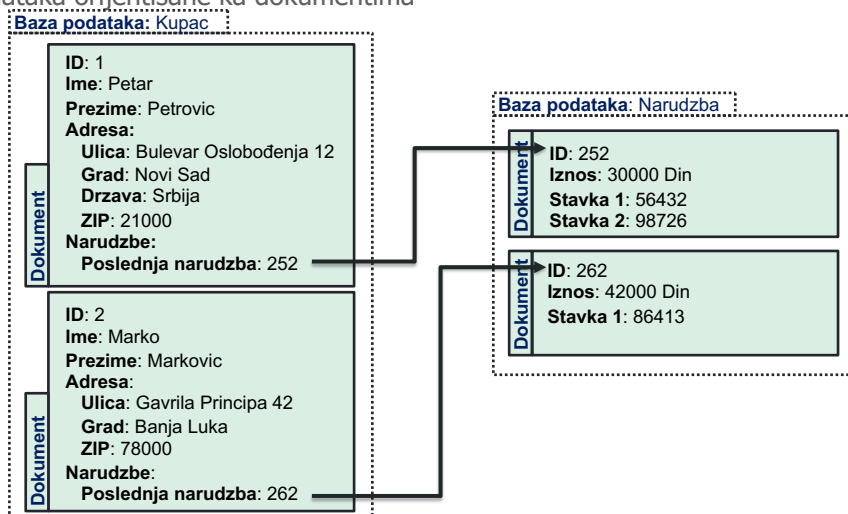
- Baze podataka orijentisane ka dokumentima
 - osobine
 - objekti mogu biti snimljeni kao dokumenti
 - serijalizacija objekta u dokument
 - jednostavna realokacija dokumenata
 - dokumenti mogu biti složeni
 - dokumenti su nezavisni
 - zapisi o jednom entitetu se nalaze u jednom dokumentu
 - jednostavni formati dokumenata
 - JSON ili XML
 - nepostojanje fiksne šeme baze podataka
 - najčešće
 - ugrađeno praćenje verzije dokumenata

42

42

NoSQL baze podataka

- Baze podataka orijentisane ka dokumentima



43

43

NoSQL baze podataka

- Baze podataka orijentisane ka dokumentima

Ključ-vrednost	CouchDB	MongoDB	RavenDB
Napisana u	Erlang	C++	C#
Platforme	Linux	Linux, Mac, Windows	Windows
Protokol	HTTP/REST	sopstveni TCP/IP	HTTP/REST
Način čuvanja podataka	disk	disk	memorija i disk
Replikacija	Peer-based	Master-Slave	plug-in
Particionisanje	da	da	da
Indeksi	da	da	da
MapReduce	da	da	da
Transakcije	ne	ne	da

44

44

NoSQL baze podataka

- Baze podataka orijentisane ka grafovima
 - engl. *graph databases*
 - zasnovane na teoriji grafova
 - odnosi između kolekcija objekata
 - grafovi se sastoje od
 - **čvorova**
 - koji predstavljaju entitete
 - **grana**
 - koje povezuju čvorove
 - **svojtava**
 - predstavljaju atribute
 - reprezentuju i čuvaju informacije

45

45

NoSQL baze podataka

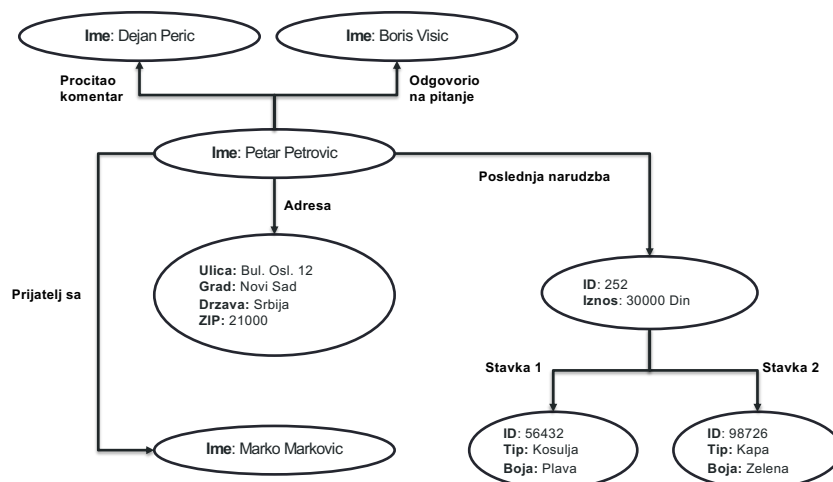
- Baze podataka orijentisane ka grafovima
 - nastale sa popularizacijom društvenih mreža
 - veliki broj korisnika koji poseduju veze sa drugim korisnicima, statusima, komentarima itd.
 - pogodne za podatke kojima rukuje semantički *web*

46

46

NoSQL baze podataka

- Baze podataka orijentisane ka grafovima



47

47

NoSQL baze podataka

- Baze podataka orijentisane ka grafovima

Ključ-vrednost	Neo4J	InfoGrid	Bigdata
Napisana u	Java	Java	Java
Platforme	bilo koja Java	bilo koja Java	bilo koja Java
Protokol	HTTP	HTTP	HTTP
Način čuvanja podataka	disk	disk, memorija i nekoliko BP	disk
Replikacija	ne	ne	da
Particionisanje	ne	ne	da
Indeksi	ne	ne	da

48

48

NoSQL baze podataka

- Uredne karakteristike i performanse NoSQL sistema i RSubP-a

	performanse	skalabilnost	fleksibilnost	kompleksnost	big data
key-value	visoke	visoka	visoka	nema	da
wide columns	visoke	visoka	srednja	mala	da
document	visoke	promenljiva	visoka	mala	da/ne
graph	promenljive	promenljiva	visoka	visoka	da/ne
RSubP	promenljive	promenljiva	visoka	srednja	ne

49

49

Reference

- Tiwari S, „*Professional NoSQL*“, John Wiley & Sons, Inc., SAD, 2011
- Todić B, „*Primena specijalizovanih baza podataka u oblasti upravljanja dokumentima*“, Master rad, FTN, 2012.
- Elmasri R, Navathe S B, „*Fundamentals of Database Systems*“, Šesto izdanje, Addison-Wesley, SAD, 2011
 - poglavlje 26