универзитет у бањој луци

електротехнички факултет

**Александар Бјелошевић**

ДОКУМЕНТ БАЗИРАНЕ БАЗЕ ПОДАТАКА

**дипломски рад**

**Бања Лука, септембар 2015.**

**Тема: ДОКУМЕНТ БАЗИРАНЕ БАЗЕ ПОДАТАКА**

**Кључне ријечи:**

**NоSQL сервери базе података**

**Документ базиране базе података**

**Комисија: проф. др Ратко Дејановић, предсједник**

**проф. др Славко Марић, ментор**

**мр Михајло Савић, члан**

**Кандидат:**

**Александар Бјелошевић**

УНИВЕРЗИТЕТ У БАЊОЈ ЛУЦИ

ЕЛЕКТРОТЕХНИЧКИ ФАКУЛТЕТ

КАТЕДРА ЗА РАЧУНАРСКУ ТЕХНИКУ

Предмет: БАЗЕ ПОДАТАКА

Тема: ДОКУМЕНТ БАЗИРАНЕ БАЗЕ ПОДАТАКА

Задатак: Кратак преглед врста NоSQL сервера база података. Детаљан опис документ базираних база података уз

анализу предности и ограничења у односу на SQL

базе података. За практичан рад реализовати веб апликацију употребом документ базираних база података.

Ментор: проф. др Славко Марић

Кандидат: Александар Бјелошевић (45/07)

Бања Лука, септембар 2015.

1. УВОД 1
2. ПРЕГЛЕД NоSQL СЕРВЕРА БАЗА ПОДАТАКА 2

2.1 Историјат NоSQL сервера база података 2

2.2 Типови NоSQL сервера база података 4

2.3 Предности и ограничења NоSQL у односу на SQL базе података 7

1. ДОКУМЕНТ БАЗИРАНЕ БАЗЕ ПОДАТАКА 8

3.1 Преглед документ базираних база података 8

3.2 MongoDB модел података 12

3.3 MongoDB агрегација 14

3.4 Сигурност у MongoDB базама података 16

3.5 MongoDB операције индексирања 18

3.6 Операција репликације у MongoDB базама података 22

3.7 MongoDB *sharding* 26

1. ПРАКТИЧАН РАД 27

5.1 MongoDB CRUD операције 27

5.2 Практична апликација 37

5.3 Дискусија 45

1. ЗАКЉУЧАК 46
2. ЛИТЕРАТУРА 47

Уз рад је приложен CD

# 2. ПРЕГЛЕД NоSQL СЕРВЕРА БАЗА ПОДАТАКА

## 2.1 Историјат NоSQL сервера база података

## NоSQL базе података пружају механизме за складиштење и добављање података кориштењем средстава која се значајно разликују од табеларних веза које су кориштене у релационим базама података. Континуалан развој веб и мобилних апликација као и константне промјене корисничких захтјева допринјели су развоју нових рјешења за чување и одржавање података.[1]

## NоSQL сервери база података обухватају широки спектар технологија које су развијене као одговор на пораст обима података о корисницима, предметима и производима, учесталости приступа овим подацима као и потребама учинка и обраде података. Релационе базе података, са друге стране, нису пројектоване да се носе са величином и агилности са којма се сусрећу модерне апликације, нити су направљене да искористе јефтиније складиштење и обрадиву моћ хардвера која је доступна данас.[2] Структура података која је кориштена у NоSQL серверима базе података(граф, документ) незнатно се разликује од оне кориштене у релационим базама података, што ће неке операције чинити бржим у NоSQL базама података, док ће остале чинити брже у релационим базама података.

Враћајући се 20 година уназад када су архитекти и програмери апликација, за потребе складиштења и обрађивања података, имали избор различитих релационих база података. Релационе базе су представљале једину опцију за девелопере и инфраструкционе тимове. За ове апликације перформансе и њихова доступност сз биле од пресудног значаја. Нови проблеми у високој доступноски на високој скали покренули су велико компаније као *Google*, *Facebook* и *Amazon* да креирају нова ријешења и нове технологије. Једно ријешење које су ове компаније развиле и унаприједиле су не-релационе базе података које су једним именом назване NоSQL базе података.[1]

Једна од великих заблуда је да термин NоSQL значи „Nо SQL“, као и да NоSQL базе података не користе SQL(прев. Структуирани Упитни Језик) као упитни језик. Али многе NоSQL базе података користе SQL, као један од многих упитних језика. Чињеница је да NоSQL боље описује оно што није, него оно што заправо јесте, израз је и даље веома користан за опис широког спектра класа база података идеалних за рјешавање проблема којима се бавимо данас.[1]

NоSQL базе података веома добро процесуирају: величину, разноврсност и брзину протока великих података.

## 2.1 Типови NоSQL сервера база података

Постоје различити приступи за класификацију NоSQL база података, од којих сваки садржи различите категорије и поткатегорије. Због различитих приступа и преклапања тешко је да се одреди преглед не-релационих база података. Ипак, основна класификација се заснива на разлици у моделу података.

Према моделу података не-релационе базе података можемо подијелити у следеће категорије[2]:

1. документ базиране базе података (енг. *Document model databases*);

2. граф базиране базе података (енг. *Graph model databases*);

3. кључ-вриједност базиране базе података (енг. *Key-value model databases*);

4. базе оријентисане ка колонама (енг. *Column model databases*).

Документ базиране базе података, за разлику од релационих база података које податке складиште у табеле, користе документе за складиштење података. Подаци се у документе типично смјештају у JSON (енг. *JavaScript Object Notation*) формату. JSON представља текстуални стандард који служи за размјену података у формату који је разумљив и читљив од стране човјека и машине. JSON формат изведен је из JavaScript језика, но сам формат је независан од било ког конкретног програмског језика. Документи пружају интуитиван и природан начин за модел складиштење података који је блиско везан са објектно-орјентисаним програмирањем – сваки документ ефикасно представља објекат. Документи могу садржавати једно или више поља, при томе свако поље представља дефинисану вриједност као што су: текстуални подаци, датуми, бинарни подаци или низови података. За разлику од релационих база података гдје су подаци распоређени у више табела и колона повезаних страним кључевима з документ базираним базама података сви записи и њихови повезани (или сродни) подаци се чувају у једном документу. Овакав приступ моделу података поједноставлјује сам приступ подацима и, у многим случајевима, елимише потребу за скупим JOIN операцијама и другим комплексним упитима.[2]

Граф базиране базе података користе графовске структуре заједно са чворовима, гранама и својствима за предстабљање и чување података. У суштини, подаци су моделовани као мрежа односа између специфичних елемената. Граф базиране базе података могу се чинити контра-интуитивне што може бити корисно за специфичне класе упита. Из дефиниције, граф базе података представља сваки систем код којега сви елементи садрже директан показивач на сусједни елемент и при чени није потребан никакав преглед индекса. Слабост граф базираних база података огледа се у томе што оне обично захтијевају да се сви подаци на једном мјесту (једном серверу) ограничајући тиме њихову скалабилност.[1]

Кључ-вриједност базиране базе података имају најједноставнији модел података међу NоSQL базама података. Оне користе претраживи индекс кључ који је повезан са вриједношћу. Ове базе података користе упарене вриједности кључева са вриједностима. Аналогију можемо повући са фајл системом, путања представља кључ док вриједност на путањи (сам фајл) представља вриједност. Релационе кључ-бриједност базе података постоје већ дуги низ година, али новије кључ-вриједност базе података спадају у NоSQL категорију база података јер су намјенски направљене да би подржале велике брзине и скалабилност чиме су одређене функционалности ускраћене (непостојање страних кључева, немогућност претраживања по вриједности и др.). Кључ-вриједност базе података се обично користе за складиштење корисничких подешавања или складиштење великих токова нетрансакционих података.[1]

Базе оријентисане ка колонама су теоретски јако слични табеларном складиштењу података у релационим базама података, једна разлика представља промјенљиви број колокна у редовима. Ове базе података углавном серијализују све вриједности одређене колоне заједно на диску, чиме све добављање специфичног атрибута у великим базама података чини веома брзо. Свака колона се може посматрати као скуп кључ-вриједност, гдје се свака колона идентификује са јединственим примарним идентификатором. Базе оријентисане ка колонама су постале познате након што је *Google* објавио *BigTable*, те је им је подстактнута популарност након *Cassandra-e* и *HBase-a.*[2]

На Табела 1.1 је табеларно представљена поређења различитих NоSQL база података:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Модел података** | **Перформансе** | **Скалабилност** | **Флексибилност** | **Комплексност** |
| Кључ-бриједност | високе | висока | висока | нема |
| Базе оријенстисане ка колонама | високе | висока | средња | мала |
| Документ базиране | високе | промјенљива (висока) | висока | мала |
| Граф базиране | промјенљиве | промјенљива | висока | висока |
| Релационе | промјенљиве | промјенљива | мала | средња |

*Табела 1.1 Поређење различитих NоSQL база података*

6. **ЛИТЕРАТУРА**

[1] IBM Coloudant, *Why NoSQL?*, <https://cloudant.com/wp-content/uploads/Why_NoSQL_IBM_Cloudant.pdf>, посјећено: 16.07.2015.године.

[2] MongoDB, *Top 5 Considerations When Evaluating NoSQL Databases*, <https://s3.amazonaws.com/info-mongodb-com/10gen_Top_5_NoSQL_Considerations.pdf>, посјећено: 17.07.2015.године.

[3] MarkLogic, *The NoSQL Generation: Embracing the Document Mode*, <http://www.marklogic.com/wp-content/uploads/2014/12/nosql-generation-embracing-document-model.pdf>, посјећено: 17.07.2015.године.