

КОМИСИЈИ ЗА СТУДИЈЕ II СТЕПЕНА ЕЛЕКТРОТЕХНИЧКОГ ФАКУЛТЕТА У БЕОГРАДУ

Комисија за студије II степена, Електротехничког факултета у Београду, на својој седници одржаној 29.08.2023. године именовало нас је у Комисију за преглед и оцену мастер рада дипл. инж. Растка Шапића под насловом „Анализа компоненте за оптимизацију упита унутар сервера базе података”. Након прегледа материјала Комисија подноси следећи

ИЗВЕШТАЈ

1. Биографски подаци кандидата

Кандидат Растко Шапић је рођен 09.04.1998. године у Приштини. Завршио је Основну школу „Филип Кљајић Фића“ у Београду као вуковац, а затим и Трећу београдску гимназију са одличним успехом. Током основне и средње школе, такмичио се из географије, историје и математике. Електротехнички факултет уписао је 2017. године. Дипломирао је на одсеку за Софтверско инжењерство 2021. године са просечном оценом 9,4. Дипломски рад одбранио је у септембру 2021. године са оценом 10. Током основних студија био је у тиму који представља факултет на међународном студентском такмичењу “Електријада”, из области Информатика. Био је ангажован као студент демонстратор на више од десет предмета са катедре за Рачунарску технику и информатику. У четвртој години основних студија је био на пракси у Мајкрософт развојном центру у Београду, где се бавио развојем великих база података у облаку. Дипломске академске – мастер студије на Електротехничком факултету у Београду, на Модулу за софтверско инжењерство уписао је у октобру 2021. године. Положио је све предмете са просечном оценом 9.8. Од септембра 2021. године је запослен у Мајкрософт развојном центру у Београду, на позицији софтвер инжењер. Ради у тиму који се бави развојем дистрибуираних база података у облаку, које се користе за аналитику података и обучавање вештаке интелигенције.

2. Извештај о студијском истраживачком раду

Кандидат Растко Шапић је као припрему за израду мастер рада урадио истраживање релевантне литературе која се односи на области којима припада тема мастер рада. Конкретно, анализирао је постојећа решења и проблеме из области интерне имплементације сервера базе података, фундаменталним концептима коришћења на основу који се заснивају интерне имплементације, типовима производа који базирају своју имплементацију на Мајкрософт серверу базе података и индустријски стандардизованих тестова перформанси у свету база података. Одабране су следеће компоненте унутар Мајкрософт сервера базе података: Оперативни систем Мајкрософт сервера базе података, компонента за рад са складиштем података, компонента за извршавање упита и компонента за оптимизацију упита. Код сценарија коришћења, истраживање се заснивало на бази два фундаментална правца: *OLTP (Online Transaction Processing)* *OLAP (Online Analytical Processing)*. Производи који су коришћени и над којима је рађено мерење перформанси су *Microsoft SQL Server 2022* и *Azure SQL Database*. На крају, за тестирање перформанси је узет златни стандард међу тестовима перформанси за сценарије који покривају подршку приликом доношења одлука *TCP-H (Transactional Processing Performance Council - H)*. Истраживањем је утврђено да је интерна логика сервера базе података веома комплексно софтверско решење које покрива све аспекте у виду обраде и манипулације над релационим базама података. Утврђено је да познавање интерних механизма може знатно утицати на перформансе, као и на дизајнирање базе података. На крају, утврђен је хардверски утицај на перформансе сервера базе података.

3. Опис мастер рада

Мастер рад обухвата 48 страна, са укупно 43 слике, 4 табела и 19 референци. Рад садржи увод, четири поглавља и закључак (укупно шест поглавља), те списак коришћене литературе, списак скраћеница, списак слика и списак табела.

Прво поглавље представља увод у сервере база података где су описане све компоненте које један сервер садржи. Објашњен је ток извршавања приликом доласка захтева за извршавање упита, детаљи везани за оперативни систем који постоји унутар сервера базе података и компоненте за рад са складиштем података. Објашњен је значај НУМА архитектуре и интерно имплементираног распоређивача (*Scheduler*) унутар оперативног система сервера. Приказан је остатак тока извршавања, превођење и оптимизација стабла извршавања, као и компонента која гарантује исправност података чак и у случају системских отказа и грешака.

Друго поглавље обрађује анализу компоненте за извршавање упита. Оно је подељено на више потпоглавља која обрађују следеће проблематике: особине оператора, анализа плана извршавања, анализа оператора за спајање две табеле, анализа оператора за агрегацију и анализа оператора за читање података са диска. Приказан је рад сваког физичког оператора за наведене групе (нпр., *Hash Join*, *Stream Aggregate*), интерна алгоритмика, као и анализа случајева коришћења у којима се ови оператори показују као добро решење. На крају овог поглавља се обрађује анализа извршавања упита у зависности од сценарија коришћења где се главни фокус своди на анализу *OLTP* сценарија и *Data Warehouse* архитектуре која користи *OLAP* сценарије.

У трећем поглављу је урађена анализа компоненте која се бави оптимизацијом упита. У почетку се поглавље осврће на целокупни процес обраде упита, као и формат стабла са којим компонента за оптимизацију упита ради. Затим је приказана логика рада ове компоненте, као и сви њени делови. Објашњена је архитектура ове компоненте са свим њеним фазама где се приказује структура која служи за чување потенцијалних кандидата за план извршавања као и процес упрошћавања упита. У последњем делу поглавља се претежно објашњавају механизми који су главни разлог због чега ова компонента даје одличне перформансе у већини случајева. На почетку, објашњен је концепт статистика у серверу, њихов дизајн, начин процењивања броја редова над табелама које имају јако пуно редова као и начин рачунања цене извршавања сваког од кандидата. На крају овог дела се анализира механизам који служи за одабир индекса приликом формирања оптималног упита.

Четврто поглавље детаљно приказује начин на који може да се оствари оптимизација над одређеним упитом. Ово поглавље се суштински дели у два потпоглавља, где се у првом потпоглављу ради анализа перформанси применом знања стечених током анализе компоненти за извршавање и оптимизацију упита. У другом потпоглављу је фокус на мерење перформанси за различите хардверске спецификације, и анализа утицаја хардвера на перформансе сервера базе података. За ове потребе коришћен је *TPC-H* тест перформанси прихваћен широм индустрије.

4. Анализа рада са кључним резултатима

Мастер рад дипл. инж. Растка Шапића се бави проблематиком оптимизација упита унутар сервера базе података, као и механизма оптимизација са стране корисника. Овај рад проналази примену у побољшању свакодневног искуства корисника, као и у побољшању дизајна базе података.

Анализа оптимизације се протеже у три гране: анализа механизма и концепата унутар сервера базе података, примена механизма оптимизације над упитима и анализа утицаја хардвера на перформансе извршавања упита.

Основни доприноси рада су: 1) приказ и анализа рада сервера базе података 2) приказ и анализа компоненте за извршавање; 3) приказ и анализа компоненте за оптимизацију упита; 4) приказ утицаја компоненте за оптимизацију у зависности у дизајна базе података и дизајна упита за извршавање; 5) приказ и анализа утицаја хардвера на перформансе рада сервера базе података.

5. Закључак и предлог

Кандидат Растко Шапић је у свом мастер раду успешно извршио анализу компоненте за оптимизацију упита, као и осталих повезаних компонената унутар сервера базе података. На основу приказане анализе, кандидат је извео експерименте који показују како познавање компоненте за оптимизацију може утицати на перформансе сервера базе података. Такође, кандидат је урадио анализу хардверског утицаја на перформансе извршавања упита.

Кандидат је исказао самосталност и систематичност у своме поступку као и иновативне елементе у решавању проблематике овог рада.

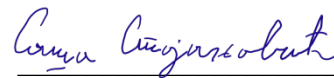
На основу изложеног, Комисија предлаже Комисији за студије II степена Електротехничког факултета у Београду да рад дипл. инж. Растка Шапића прихвати као мастер рад и кандидату одобри јавну усмену одбрану.

Београд, 07.09.2023. године

Чланови комисије:



др Милош Цветановић, ванр.проф.



др Саша Стојановић, доцент