

NASTAVNO-NAUČNOM VEĆU ELEKTROTEHNIČKOG FAKULTETA U BEOGRADU

Nastavno-naučno veće Elektrotehničkog fakulteta u Beogradu imenovalo nas je 8.7.2014. za članove Komisije za pregled i ocenu magistarske teze kandidata Darka Jovića, diplomiranog inženjera elektrotehnike, pod naslovom "Realizacija distribuirane baze podataka u P2P okruženju".

Pošto smo pregledali ovu tezu podnosimo Nastavno-naučnom veću sledeći

IZVEŠTAJ

1. Osnovni podaci i kandidatu

Darko Jović je rođen 21.12.1978. godine u Zagrebu. Osnovne studije je završio na Elektrotehničkom fakultetu u Beogradu, smer Računarska tehnika i informatika, sa prosečnom ocenom 8.26 (osam i 26 od 100), uz ocenu na diplomskom radu 10 (deset). Diplomirao je 2003. godine, uspešno odbranivši diplomski rad na temu: "Promenljivi korisnički interfejs ASP.NET aplikacija". Od 2003. radio je u nekoliko softverskih kompanija i to na projektima vezanim za elektronsko bankarstvo; sisteme za upravljanje, kategorizaciju i skladištenje digitalnih fotografija; sisteme za efikasno upravljanje keširanih sadržaja (u saradnji sa Dow Jones, USA). U tom periodu bio je i angažovan na nekoliko istraživačkih projekata iz oblasti P2P sistema (u saradnji sa Fraunhofer IPSI institutom) i velikih sistema za skladištenje podataka (u saradnji sa StorageTek, USA). Od 2006. do 2010. godine radio je kao istraživač saradnik u Inovacionom centru Elektrotehničkog fakulteta u Beogradu, gde je učestvovao na realizaciji nekoliko domaćih projekata iz oblasti poboljšanja elektronske uprave, kao i internacionalnog FP7 projekta "We-Go, Enhancing Western Balkan eGovernment expertise". Trenutno radi u kompaniji GTech kao vođa tima za razvoj multikorisničke platforme za Internet kazino igre. Poslediplomske studije na Elektrotehničkom fakultetu u Beogradu, smer Arhitektura i organizacija računarskih sistema i mreža, upisao je 2003. godine i položio sve ispite predviđene Nastavnim planom i programom magistarskih studija, sa prosečnom ocenom 10 (deset).

2. Osnovni podaci o radu

Predmet rada je realizacija distribuirane baze podataka u P2P okruženju, koja svoje operacije izvršava nad XML dokumentima, odnosno delovima dokumenata preko postojećih i široko korišćenih Document Object Model (DOM) interfejsa. Uzimanje P2P sistema za osnovu distribuirane baze podataka, obrazloženo je karakteristikama P2P sistema kao što su niski troškovi infrastrukture, skalabilnost, decentralizovanost, ad-hoc komunikacija i sl. Međutim zbog svoje dinamične prirode, implementacija potrebnih funkcionalnosti distribuirane baze podataka kao što su lociranje podataka, integritet i konzistentost, kao i njihova pretraga u okruženju gde dinamika dolaska i odlaska čvorova može biti veoma velika, nailazi na probleme kojima se rad bavi i predlaže rešenja za njihovo prevazilaženje. Predložena je arhitektura sistema koja ispunjava navedene funkcionalnosti, odnosno obezbeđuje dostupnost i konzistentnost sa visokom verovatnoćom, kao i funkcionalnost distribuirane pretrage podataka. Kao evaluacija predloženog rešenja napravljen je i prototip, čiji je cilj da pokaže da je sistem moguće sprovesti u praksi.

Rad sadrži devet glava, uključujući uvod, priloge i literaturu. Preostale glave uključuju: formulaciju problema, analizu postojećih rešenja, predloženo rešenje, implementaciju prototipa rešenja, zaključak i dalje pravce istraživanja. Rad sadrži 99 strana, 28 slika i 52 reference.

3. Analiza rada sa ključnim rezultatima

U prvoj glavi dat je sažetak rada, u kojem su u kratkim crtama navedeni problem kojim se rad bavi, postojeća rešenja iz date oblasti, kao i predloženo rešenje i njegov doprinos.

U drugoj glavi rada dat je uvod u oblast. Opisane su osnovne osobine distribuiranih sistema, kao i njihova klasifikacija. Zatim su opisani P2P sistemi kao podskup distribuiranih sistema, date su njihove osobine i razlozi zbog kojih se koriste. Zatim je dat opis distribuiranih heš tabela i primer Pastry rutirajućeg algoritma pomoću kojeg se deterministički pronalazi određište poruke. Pored ovoga, dat je kratak opis JxTa specifikacije. Na kraju uvoda date su osnove XML tehnologije, uključujući opis XPath upitnog jezika koji se koristi za pretragu delova XML dokumenta.

U trećoj glavi data je formulacija problema. Kao glavni cilj rada navedena je implementacija distribuiranog sistema za manipulaciju dokumenata u P2P okruženju pri čemu sistem treba da zadovolji niz uslova kao što su: deterministička lokacija podataka, garantovani nivo dostupnosti i konzistentnosti podataka, balansiranost, granulacija podataka, transparentna distribuiranost podataka, funkcionalnost pretrage, standardizovani aplikativni interfejs i zadovoljavajuće performanse.

U četvrtoj glavi je data analiza postojećih rešenja iz oblasti upravljanja podataka u distribuiranim P2P sistemima. Razmatrana su dva ključna problema distribuiranih sistema: dostupnost i konzistentnost podataka. Navedene su postojeće tehnike koje se koriste za povećanje nivoa dostupnosti i konzistentnosti podataka u distribuiranim sistemima, kao i implikacije njihove potencijalne primene u P2P sistemima. Zatim je dat pregled osobina i nedostataka postojećih P2P sistema za čuvanje, razmenu i arhiviranje podataka kao što su Napster, Gnutella, KaZaA, eDonkey, BitTorrent, FreeNet, CFS, PAST, OceanStore, Ivy i dr. Na kraju glave dat je njihov uporedni tabelarni prikaz po karakteristikama kao što su: korišćeni rutirajući protokol, determinističko lociranje podataka, skalabilnost, korišćenje replikacije, keširanje, mogućnost pretrage, mogućnost izmene podataka i sl.

U petoj glavi je opisano predloženo rešenje koje podrazumeva implementaciju distribuirane baze podataka koja manipuliše XML dokumentima u P2P okruženju. Distribucija dokumenata u sistemu je takva da nijedan čvor ne sadrži ceo dokument, već su delovi dokumenta distribuirani na različite čvorove u mreži. Dokumenti se kreiraju i modifikuju u realnom vremenu i svaki čvor im pristupa na uniforman način, odnosno ne zna ništa o distribuciji dokumenta u sistemu. Kao osnova sistema uzeta je DHT implementacija sa Pastry rutirajućim algoritmom, pri čemu se za generisanje identifikatora, koji ujedno određuju i adresu podatka u mreži, koristi heš funkcija kao što je MD5 ili SHA-1. Dobra raspodela identifikatora u prostoru obezbeđuje balansiranost sistema, jer će generisani identifikatori biti dovoljno udaljeni međusobno tako da će biti smešteni na različitim čvorovima sistema. Dostupnost podataka je obezbeđena replikacijom, pri čemu je implementiran implicitni registar replika. Registar podrazumeva jednostavan algoritam generisanja identifikatora replika na osnovu identifikatora originalnog podatka, tako da svaki čvor zna lokaciju svake replike u sistemu. Primenjen je model eventualne konzistentnosti podataka, implementiran kao kombinacija verzionisanja objekata i dvofaznog komit protokola. Sistem pruža svoju funkcionalnost preko Document Object Model (DOM) interfejsa, a zaokružen je funkcionalnošću distribuirane XPath pretrage. Ovde su date i uvedene pretpostavke pod kojim sistem ispunjava traženu funkcionalnost. U nastavku je opisana arhitektura sistema, sa slikom i opisom svake od komponenata. Zatim su opisane korisničke funkcionalnosti sistema i detaljan opis rešenja problema determinističke lokacije objekata, skalabilnosti i balansiranosti, dostupnosti, konzistentnosti i distribuirane pretrage. Na kraju je data i analiza očekivanih performansi sistema.

U šestoj glavi opisana je implementacija prototipa sistema u programskom jeziku Java. Dat je opis glavnih implementacionih klasa, a njihova zavisnost i struktura glavnih paketa predstavljeni su UML dijagramima klasa. Opisana je implementacija svake od korisničkih funkcionalnosti sistema korišćenjem UML dijagrama sekvenci.

U sedmoj glavi dat je zaključak, dok su u osmoj glavi navedeni pravci mogućih unapređenja sistema. I konačno, deveta glava sadrži spisak referenci.

4. Zaključak i predlog

Kandidat Darko Jović je u svojoj magistarskoj tezi, na osnovu analize postojećih rešenja iz oblasti distribuiranih sistema za skladištenje podataka, identifikovao skup problema sa kojima se takvi sistemi suočavaju. P2P mreže kao veoma dinamično okruženje donose izazove na polju zadovoljenja konzistentnosti

i dostupnosti podataka. Osnovni problem na koji je kandidat ukazao se odnosi na realizaciju distribuirane baze podataka u P2P okruženju koja treba da bude decentralizovana, skalabilna, otporna na otkaze i da obezbeđuje dostupnost podataka na visokom nivou u uslovima gde postoji veći broj čvorova u mreži. Kandidat je uradio analizu postojećih P2P sistema koji se bave skladištenjem i upravljanjem podacima i identifikovao njihove nedostatke: nisu potpuno decentralizovani, niska granularnost podataka, nemogućnost izmene upisanog podatka i dr. Prikazan je sistematičan pristup analizi postojećih rešenja na osnovu koje je kandidat predložio jedno sveobuhvatno rešenje koje se zasniva na distribuciji delova XML dokumenta u P2P mreži koristeći Pastry DHT implementaciju kao osnovu. Sistem u potpunosti ispunjava uslove zadatog problema, u skladu sa uvedenim pretpostavkama, a to je i pokazano implementacijom prototipa sistema. Kandidat je i identifikovao određene nedostatke datog rešenja i dao pravce budućeg istraživanja za njihovo otklanjanje.

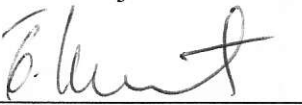
Na osnovu izloženog, članovi Komisije predlažu Naučno-nastavnom veću Elektrotehničkog fakulteta u Beogradu da prihvati magistarski rad kandidata Darka Jovića, diplomiranog inženjera elektrotehnike, pod naslovom "Realizacija distribuirane baze podataka u P2P okruženju" i odobri njegovu javnu odbranu.

U Beogradu, 21.08.2014. godine


Članovi Komisije:



prof. dr Veljko Milutinović



prof. dr Boško Nikolić



prof. dr Predrag Jančić