

MOGUĆNOSTI PRIMENE SPOT PRICING MEHANIZMA U CLOUD OKRUŽENJU

Branka Mikavica, Vesna Radonjić Đogatović, Aleksandra Kostić-Ljubisavljević
Univerzitet u Beogradu - Saobraćajni fakultet

Sadržaj: *Intenzivan razvoj cloud computing-a dovodi do neophodnosti unapređenja adekvatnih tarifnih mehanizama. Najširu primenu još uvek imaju fiksni tarifni mehanizmi koji se odlikuju fiksnim cenama u jedinici vremena. Kako bi se iskoristile prednosti cloud computing-a neophodno je primeniti dinamičke tarifne mehanizme koji bolje oslikavaju odnos ponude i tražnje. U ovom radu je analizirana mogućnost primene spot pricing mehanizma tarifiranja resursa cloud provajdera. Prikazane su osnovne karakteristike spot pricing mehanizma i aukcija kao načina realizacije. Razmatrane su prednosti i nedostaci ovog tarifnog mehanizma kako sa aspekta cloud provajdera, tako i sa aspekta korisnika cloud servisa.*

Ključne reči: *spot pricing, cloud computing, cloud provajderi, cloud servisi, korisnici cloud servisa*

1. Uvod

Pojava *cloud computing*-a predstavlja fundamentalnu promenu u načinu na koji se informacione tehnologije razvijaju, implementiraju, inoviraju, održavaju i tarifiraju. Primena *cloud computing*-a obezbeđuje brojne prednosti svim korisnicima: jednostavnija instalacija softvera, održavanje i centralizacija; skladištenje podataka i korišćenje *cloud* infrastrukture; ne postoji problem obezbeđivanja hardverskih resursa zbog skoro beskonačnih *cloud* resursa dostupnih na zahtev; postoji mogućnost plaćanja samo za one resurse koji su korišćeni, bez sklapanja dugoročnih ugovora (na primer, po času ili danu korišćenja) sa mogućnošću oslobađanja resursa nakon završetka sesije. Intenzivan razvoj i primena *cloud computing*-a uslovljava sve veći značaj razvoja tarifnih mehanizama za naplatu korišćenja računarskih resursa. Imajući u vidu da se računarski resursi koje *cloud* provajderi iznajmljuju korisnicima mogu smatrati privremenim resursima, kao i to da tražnja za korišćenje računarskih resursa nije ravnomerna tokom vremena, neophodno je razvijati dinamičke koncepte tarifiranja u cilju maksimiziranja prihoda [1]. Tržišno orijentisani tarifni mehanizmi koji se zasnivaju na prikupljanju ponuda od korisnika uz alokaciju resursa, pogodni su za realizaciju dinamičkih oblika tarifiranja. Ovakvi tarifni mehanizmi dobijaju sve veći značaj pri iznajmljivanju neiskorišćenih kapaciteta *cloud* provajdera [2-3]. Pokazalo se da mehanizmi aukcija mogu biti efikasni za realizaciju dinamičkog tarifiranja *cloud* resursa obzirom da: podstiču korisnike da preispitaju pravu vrednost zahtevanih resursa (tj. nude cenu koju su spremni da plate za određene resurse), obezbeđuju da se resursi alociraju onim korisnicima koji ih najviše vrednuju, i korektno

postavljaju cene resursa u skladu sa uslovima ponude i tražnje stvaranjem konkurencije između korisnika [4]. U literaturi se može pronaći sledeća klasifikacija tarifnih mehanizama u *cloud computing*-u: *Pay-per-Use*, *Subscription Based*, i *Hybrid pricing model* [5]. Ipak, različiti provajderi različito su klasifikovali ove tarifne mehanizme. Tako Amazon primenjuje tarifiranje za: *reserved*, *on-demand* i *spot* instance. Slično tome, Microsoft Azure primenjuje *Pay-as-you-go* i mesečne tarifne mehanizme. Rezervacija kao tarifni mehanizam omogućava rezervisanje resursa unapred u određenom vremenskom intervalu uz definisanje SLA (*Service Level Agreement*) sa *cloud* provajderom. Na taj način, resursi su dodeljeni različitim korisnicima na određeno vreme i na raspolaganju su im u bilo kom trenutku, bez obzira da li se dati resursi koriste ili ne. *Pay-per-Use* tarifni mehanizam podrazumeva da se resursi dodeljuju na zahtev a tarifiranje tih resursa se vrši u zavisnosti od korišćenja u jedinici vremena. Hibridni mehanizam tarifiranja predstavlja prelazni mehanizam između tarifiranja u zavisnosti od korišćenja i rezervacije. Rezervisani resursi moraju biti obezbeđeni unapred za određen vremenski interval. Dodatni resursi mogu se obezbediti na zahtev i biti naplaćeni u zavisnosti od korišćenja resursa. Najnoviji mehanizam tarifiranja u *cloud computing*-u je *spot pricing* koji za cilj ima bolje iskorišćenje resursa *cloud* provajdera, naročito u periodima nižeg opterećenja. Zasniva se na aukcijama i obezbeđuje niže cene u odnosu na *on-demand* i mehanizam rezervacije. U slučaju primene ovog tarifnog mehanizma, raspoloživost resursa predstavlja veliki izazov.

Rad je koncipiran na sledeći način. Nakon uvodnog razmatranja sledi pregled osnovnih pojmova i definicija vezanih za *cloud computing*, klasifikacija učesnika u procesu obezbeđivanja *cloud* servisa i pregled postojećih kategorija *cloud* servisa. Osnovne karakteristike *spot pricing* mehanizma kao predstavnika dinamičkih mehanizama tarifiranja u *cloud computing*-u prikazane su u trećem poglavlju. Prednosti i nedostaci ovog tarifnog mehanizma sa aspekta *cloud* provajdera i korisnika *cloud* servisa analizirane su u četvrtom poglavlju. Poslednje poglavlje rada predstavlja zaključna razmatranja.

2. Cloud computing - osnovni pojmovi i definicije

Prema ITU (*International Telecommunication Union*), *cloud computing* predstavlja paradigmu za omogućavanje pristupa skalabilnim, elastičnim i deljivim fizičkim ili virtuelnim resursima uz obezbeđivanje i administiranje servisa na zahtev [6]. U kontekstu *cloud computing*-a, često je potrebno diferencirati zahteve i aktivnosti određenih učesnika. Učesnici su entiteti određeni svojim funkcijama (ulogama) koje predstavljaju skup aktivnosti. Učesnik može imati više od jedne funkcije u bilo kom trenutku i može se ograničiti na samo određeni podskup aktivnosti za datu funkciju. Učesnici u *cloud computing*-u su korisnik *cloud* servisa, *cloud* partner i *cloud* provajder. Korisnik *cloud* servisa je učesnik koji je u poslovnoj relaciji sa *cloud* partnerom ili *cloud* provajderom u cilju korišćenja *cloud* servisa. *Cloud* partner je učesnik koji obezbeđuje podršku aktivnostima *cloud* provajdera, korisnika, ili i jedno i drugo. Primer *cloud* partnera je *cloud* broker ili *cloud* revizor. *Cloud* provajder je učesnik koji obezbeđuje *cloud* servise. Pored osnovne funkcije, *cloud* provajder je zadužen za monitoring i održavanje, a može imati i dodatne funkcije kao što su *business manager*, *service manager*, *network provider*, *security and risk manager* itd [6].

Osnovne karakteristike *cloud computing*-a su pristup mreži bilo kada i bilo gde, mogućnost monitoringa servisa, mogućnost iznajmljivanja resursa, velika elastičnost i skalabilnost [6]. Jedno od osnovnih karakteristika *cloud computing*-a je svojstvo dostupnosti fizičkih i virtuelnih resursa kojima se može pristupiti standardnim mehanizmima sa različitih korisničkih platformi. Na taj način, *cloud computing* obezbeđuje dodatne pogodnosti tako da korisnici mogu pristupiti fizičkim i virtuelnim resursima sa bilo koje lokacije, sa bilo kog korisničkog uređaja dokle god imaju pristup mreži. Važna osobina *cloud computing*-a je da se obezbeđivanje *cloud* servisa može meriti, nadgledati, kontrolisati i tarifirati. Na osnovu toga korisnici mogu platiti samo one resurse kojima je pristupano. Opsluživanje na zahtev podrazumeva da korisnici *cloud* servisa mogu koristiti računarske kapacitete, po potrebi, automatski ili sa minimalnom interakcijom sa *cloud* provajderom. Na taj način *cloud computing* obezbeđuje relativno smanjenje troškova, vremena i resursa. Sa aspekta korisnika, fizički i virtuelni *cloud* resursi su praktično neograničeni i mogu se iznajmiti različiti kapaciteti u bilo koje vreme automatski, u zavisnosti od ograničenja SLA, čime se značajno pojednostavljuje proces planiranja kapaciteta.

U zavisnosti od resursa koji se zauzimaju, postoje tri različite funkcionalnosti koje omogućava *cloud computing* a koje se odlikuju minimalnim međusobnim podudaranjem [6]:

- *application capabilities type* (korisnik pristupa aplikacijama koje obezbeđuje *cloud* provajder)
- *infrastructure capabilities type* (korisnik ima pravo pristupa resursima za obradu i skladištenje podataka ili resursima mreže)
- *platform capabilities type* (korisnik može pristupati i upravljati aplikacijama korišćenjem jednog ili više programskih jezika i jednog ili više izvršnih okruženja koje obezbeđuje *cloud* provajder).

Ove funkcionalnosti treba razlikovati od kategorija *cloud* servisa. Kategorija *cloud* servisa predstavlja grupu *cloud* servisa koja poseduje neki zajednički skup karakteristika i može uključivati jednu ili više vrsta funkcionalnosti. U nastavku je navedeno i opisano nekoliko kategorija *cloud* servisa.

CaaS (*Communication as a Service*) je kategorija *cloud* servisa u kojoj se korisnicima *cloud* servisa kao funkcionalnost obezbeđuje interakcija u realnom vremenu. CompaaS (*Compute as a Service*) je kategorija *cloud* servisa u kojoj se korisnicima obezbeđuje pristup i korišćenje računarskih resursa neophodnih za izvršavanje aplikacija. DSaaS (*Data Storage as a Service*) predstavlja kategoriju *cloud* servisa koja podrazumeva obezbeđivanje resursa i pripadajućih funkcionalnosti za skladištenje podataka. IaaS (*Infrastructure as a Service*) je kategorija *cloud* servisa u kojoj se kao funkcionalnost obezbeđuje *cloud* infrastruktura. Naas (*Network as a Service*) korisnicima obezbeđuje konektivnost i pripadajuće mrežne funkcionalnosti. Paas (*Platform as a Service*) je kategorija *cloud* servisa u kojoj je funkcionalnost koja se obezbeđuje korisnicima platformskog tipa (*platform capabilities type*). SaaS (*Software as a Service*) predstavlja kategoriju *cloud* servisa koja obezbeđuje aplikacijske funkcionalnosti korisnicima.

Prethodno navedene kategorije *cloud* servisa sistematično su prikazane u Tabeli 1 u zavisnosti od funkcionalnosti koje podržavaju.

Tabela 1. Kategorije cloud servisa i vrste cloud funkcionalnosti [6]

Kategorija cloud servisa	Vrsta cloud funkcionalnosti		
	Infrastruktura	Platforma	Aplikacije
CompaaS	X		
CaaS		X	X
DSaaS	X	X	X
IaaS	X		
NaaS	X	X	X
PaaS		X	
SaaS			X

U zavisnosti od organizacije, kontrole i deljenja fizičkih i virtuelnih resursa mogu se razlikovati nekoliko modela implementacije *cloud computing*-a: javni, privatni, *community* i hibridni. Model implementacije koji podrazumeva da je *cloud* servise moguće obezbediti bilo kojem korisniku, pri čemu resurse kontroliše *cloud* provajder predstavlja javni *cloud*. Javni *cloud* može biti u vlasništvu, odnosno, njime se može upravljati i može se eksploatisati od strane poslovne, akademske ili državne organizacije, ili njihove kombinacije. Javno *cloud* okruženje je takvo da korisnici *cloud* servisa imaju malo ili nemaju nikakva ograničenja. Dostupnost *cloud* servisa korisnicima može biti predmet regulative. Privatni *cloud* podrazumeva da se *cloud* servisi obezbeđuju na osnovu ekskluzivnog prava jednom korisniku koji je ujedno zadužen i za kontrolu resursa. Privatni *cloud* može biti u vlasništvu, odnosno, njime se može upravljati i može se eksploatisati od strane jedne organizacije, ili treće strane. Korisnik *cloud* servisa takođe može ovladati pristup drugim učesnicima kako bi uvećali svoju dobit. U slučaju privatnih *cloud* modela, postoje stroga ograničenja korisnika tako da mogu pripadati samo jednoj organizaciji. *Community cloud* model implementacije podrazumeva da se *cloud* servisi obezbeđuju na ekskluzivnoj osnovi i mogu se deliti određenom skupu korisnika koji imaju iste zahteve i međusobno su povezani, pri čemu se resursi kontrolišu od strane najmanje jednog člana skupa. *Community cloud* može biti u vlasništvu, odnosno, njime se može upravljati i može se eksploatisati od strane jedne ili više organizacija, treće strane, ili neke kombinacije prethodno navedenog. Ograničenja za korisnike su manje restriktivna u odnosu na privatne *cloud* modele. Hibridni *cloud* model primenjuje najmanje dva prethodno opisana modela implementacije. Implementirani modeli ostaju jedinstveni entiteti ali predstavljaju deo hibridnog *cloud*-a tako da je moguća njihova međusobna interoperabilnost i prenostivost podataka i aplikacija. Hibridni *cloud* može biti u vlasništvu, odnosno, njime se može upravljati i može se eksploatisati od strane date organizacije ili treće strane. Ovaj model predstavlja situaciju u kojoj su neophodne interakcije između dva različita modela implementacije, a ograničenja su definisana samim modelima koji učestvuju u formiranju hibridnog *cloud*-a. Kako bi se iskoristile prednosti *cloud computing*-a, a pre svega skalabilnost na zahtev, *cloud* provajderi obezbeđuju računarske resurse u cilju opsluživanja vršnih vrednosti tražnje. Nivo iskorišćenosti ovih resursa se povećava deljenjem tih resursa različitim korisnicima. Razlog za to je taj da se vršne vrednosti tražnje različitih korisnika dešavaju u različitim vremenskim intervalima koji se ne moraju preklapati. Međutim, veliki deo računarskih resursa biva neiskorišćen u periodima vanvršnog opterećenja. Kako bi se povećala iskorišćenost ovih resursa, *cloud* provajderi uvode *spot* instance i dinamičko tarifiranje [7].

3. Osnovne karakteristike *spot pricing* tarifnog mehanizma

Kako bi se mogao garantovati uspeh i održivost *cloud* okruženja, neophodni su razvoj i implementacija prikladnih tarifnih mehanizama. Danas, fiksno tarifiranje je dominantni tarifni mehanizam [8]. U okviru toga, najčešće se primenjuje *on-demand*, gde se servis tarifira po fiksnoj ceni u jedinici vremena. Međutim, iako je tražnja sama po sebi varijabilna, može se uočiti da preostaje jedan deo neiskorišćenih resursa *cloud* provajdera. U cilju boljeg iskorišćenja tih neupotrebljenih resursa predložen je *spot pricing* koji obezbeđuje niže jedinične cene za *spot* instance čime se podstiče rast tražnje ali uz veću toleranciju na kašnjenje i otkaze. Činjenica je da su *spot* cene u opštem slučaju dosta niže od *on-demand* cena. Iz tog razloga se *spot pricing* smatra troškovno najefikasnijim tarifnim mehanizmom sa aspekta korisnika. Međutim, još uvek postoji veliko nepoverenje kada se radi o usvajanju ovog tarifnog mehanizma pa i *cloud* provajderi kao i korisnici oklevaju pri njegovom izboru. Ogromna većina *cloud* provajdera ne primenjuje *spot pricing*. Razlog za to leži upravo u nedostacima *spot pricing*-a. Za razliku od statičkih tarifnih mehanizama kao što su *on-demand* i rezervacija, *spot pricing* može biti vrlo složen sa aspekta implementacije. Pored toga, i ponuda i tražnja *spot* resursa su promenljive veličine, korisnici mogu biti izloženi neregularnim varijacijama u ceni za obezbeđivanje servisa i otkazima pri obezbeđivanju servisa.

Još uvek vladaju oprečna mišljenja u naučnoj zajednici kada se radi o *spot pricing*-u. Kao što je rečeno, obezbeđivanje *spot* instanci se čini kao efikasan pristup za privlačenje novih korisnika, bolje iskorišćavanje raspoloživih resursa i uvećanje prihoda. Međutim, može se pokazati da u nekim slučajevima fiksno tarifiranje obezbeđuje više prihode [9]. Ipak, teško je dati pravu procenu, obzirom da je teško doći do potrebnih informacija izuzev o kratkoročnoj promeni *spot* cena. Iako su fiksni tarifni mehanizmi još uvek najčešće korišćeni, *spot pricing* se preporučuje u cilju formiranja boljih tržišnih uslova u *cloud* okruženju. Poželjno je izbegavati primenu fiksnog tarifiranja kada god je to moguće. *Spot pricing* kao dinamički mehanizam se već dosta koristi od strane aviokompanija i elektrodistribucija. Međutim, uzimajući u obzir karakteristike *cloud computing*-a, nemoguće je jednostavno preslikati iskustva iz drugih industrija. *Spot* instance predstavljaju one resurse koji preostaju *cloud* provajderu nakon zauzimanja svih rezervisanih i *on-demand* instanci. Ove instance se tarifiraju po *spot* cenama koje definiše *cloud* provajder i koje se periodično menjaju u zavisnosti od ponude i tražnje za datim kapacitetima. Korisnik *cloud* resursa na određen način vrednuje obezbeđivanje svojih zahteva za pristup servisu koje se izvršava na određenom tipu *spot* instance. Postojeći *spot pricing* tarifni mehanizmi rade po principu aukcija. Po tom principu, *cloud* resursi se iznajmljuju po takvoj tržišnoj ceni da je ukupna tražnja jednaka ukupnoj ponudi. Svaka cena koja se nalazi u rasponu od najviše ponude, koja nije dobila u aukciji, i najniže ponude, koja jeste dobila u aukciji, izjednačava ponudu i tražnju. U ovom slučaju, najniža ponuda koja dobija u aukciji je upravo *spot* cena. Nakon dostavljanja korisničke aukcijske ponude, platforma za aukciju određuje tržišnu cenu, koja se naziva *spot* cena. Ta cena se definiše na osnovu aukcijskih ponuda različitih korisnika (tj. u zavisnosti od tražnje) i raspoloživih resursa (ponuda). Korisnici čija aukcijska ponuda premašuje *spot* cenu iznajmljuju *spot* instancu. Preostali korisnici se ne opslužuju. Korisnici mogu definisati svoju aukcijsku ponudu u bilo kom trenutku neograničen broj puta. Međutim, iznos koji korisnici plaćaju za iznajmljivanje *spot* instance za bilo koje vreme opsluge

zavisi od *spot* cene, a ne od korisničke aukcijske ponude. Pristup *spot* instancama i optimizacija troškova korisnika zavisi od korisničke aukcijske ponude. Poznavanje optimalne aukcijske ponude bi pomoglo korisnicima da obezbede željene resurse po nižoj ceni kroz aukcijsku platformu. Stoga, procena optimalne aukcijske ponude u bilo kom trenutku je od velikog značaja.

Cloud provajder Amazon usvojio je aukciju kao pristup za proširenje svojih tarifnih mehanizama *spot* instancama za *Amazon Elastic Compute Cloud (EC2)*. U ovom slučaju, korisnici dostavljaju svoje ponude za iznajmljivanje resursa na jedan čas. Na osnovu toga, *cloud* provajder uspostavlja tržišno orijentisanu *spot* cenu po kojoj se naplaćuje korišćenje resursa. Istovremeno, vrši se obustavljanje iznajmljivanja koje se u datom momentu naplaćuje po nižoj ceni od te tržišno uspostavljene *spot* cene. Iako Amazon nije jedini provajder koji primenjuje dinamičko tarifiranje, trenutno je jedini provajder koji javno obezbeđuje aukciju kao mehanizam za iznajmljivanje resursa. Pokušaja da se to primeni u praksi je bilo i ranije [10]. Ipak, Amazon ne otkriva detaljne informacije u vezi sa mehanizmom aukcije i izračunavanjem *spot* cena. U ovom trenutku, razvoj efikasnog i fer mehanizma aukcije koji maksimizira profit predstavlja izazov za istraživanje koji je od velikog značaja za *cloud* provajdere. Neiskorišćeni kapaciteti *cloud* provajdera koji se mogu iznajmiti na *spot* tržištu su najčešće daleko veći od tražnje. Stoga, provajder potencijalno može prihvatiti sve zahteve korisnika za iznajmljivanje resursa. U tom kontekstu, poznati mehanizmi aukcija, kao što je *second-price Vickrey* aukcija, mogu biti neadekvatni za ostvarivanje očekivanih profita provajdera [4]. U opštem slučaju, kada je ponuda veća od tražnje, ponuđači su manje motivisani da podstiču konkurenciju, što može uticati na to da provajderi ne mogu obezbediti optimalan prihod. Iz tog razloga provajderi pokušavaju uspostaviti takav mehanizam aukcije koji bi maksimizirao prihod uz podsticanje korisnika da ponude realne vrednosti za iznajmljivanje resursa *cloud* provajdera. Mehanizam aukcije je istinit ako za svakog ponuđača i i svaki izbor vrednosti svih ostalih ponuđača, dominantna strategija ponuđača i je da verno prikaže privatne informacije. Strategija je dominantna ako ponuđač ne može uvećati korisnost ostvarenu učestvovanjem u aukciji u odnosu na korisnost koju bi ostvario izlaskom iz aukcije. Takav mehanizam se može formirati pod uslovom da su dostupne sve informacije o vrednostima ponuda. Međutim, to često nije slučaj, tako da tarifiranje u mnogome zavisi od preciznosti analize tržišta. Takva analiza se mora redovno ažurirati u cilju prilagođavanja promenama na tržištu. Štaviše, obzirom da su korisnici *cloud* servisa locirani širom sveta a isti servis se može obezbeđivati sa različitim kašnjenjem, pretpostavka da su vrednovanja svih ponuđača/korisnika nezavisna i da imaju istu raspodelu ne važi. Pored toga, nemoguće je da tarifni mehanizam istovremeno poseduje sledeća svojstva: istinitost, efikasnost, racionalnost sa aspekta svakog korisnika sa optimalnim troškovima.

4. Prednosti i nedostaci *spot pricing* tarifnog mehanizma

U literaturi se mogu naći oprečna mišljenja u pogledu *spot pricing*-a [11-13]. Dinamičko tarifiranje se principijelno smatra efikasnijim u pogledu alokacije resursa. Efikasnost iskorišćenosti resursa je maksimizirana ako se primenjuju mehanizmi aukcija za tarifiranje obzirom da se *cloud* resursi dodeljuju korisnicima koji ih najviše vrednuju [13]. Drugim rečima, *spot pricing* tarifni mehanizam pruža uvid u tržišnu vrednost *cloud* resursa, posebno kada se radi o alokaciji ograničenih resursa sa potencijalno velikom

tražnjom. U praksi, *spot pricing* je ugrađen u mehanizme alokacije resursa i maksimizacije profita. Kada se radi o alokaciji resursa, uvreženo je mišljenje da *spot pricing* povećava iskoršćenje resursa *data* centara jer podstiče iznajmljivanje *cloud* resursa po nižim cenama koji bi u suprotnom biti neiskorišćeni. Pored toga, fiksni tarifni mehanizmi ne mogu efikasno oslikati tržišne uslove, naročito kada se radi o slučajevima niske i visoke tražnje. Nasuprot tome, *spot pricing* je obećavajuća alternativa kojom se može postići bolja reakcija na nepredvidive varijacije tražnje. Izmeštanjem vršnih vrednosti tražnje, *spot pricing* može uravnotežiti neke zahteve za obezbeđivanje servisa na osnovu cenovnih podsticaja i tako dovesti do efikasnijeg korišćenja infrastrukture *cloud* resursa. Kao u slučaju maksimizacije profita, efikasnost *spot pricing*-a se može razmatrati sa aspekta uštede u troškovima ili maksimizaciji prihoda. Povećanjem iznajmljivanja neiskorišćenih resursa, primena *spot pricing* omogućava opsluživanje različitih korisnika u cilju ostvarivanja viših prihoda. Fiksni mehanizmi tarifiranja obezbeđuju zadovoljavajuće prihode [9], pa ipak, više empirijskih studija pokazuje da pod određenim uslovima *spot pricing* obezbeđuje više prihode od fiksnih mehanizama tarifiranja [13, 14-16], naročito u slučaju alokacije resursa velikom broju korisnika.

Spot pricing obezbeđuje još jednu važnu pogodnost za *cloud* provajdere, a to je cenovno regulisanje tražnje i mogućnost otkazivanja obezbeđivanja svojih resursa po potrebi, što se može smatrati mogućnošću prilagođavanja dinamičkim uslovima tražnje. Ova mogućnost kontrole pre svega štiti korisnike *cloud* resursa od monopolizacije kroz *spot pricing*. U nekim situacijama, *cloud* provajderi mogu imitirati *spot pricing* u cilju kreiranja konstantnih promena na *spot* tržištu kako bi se maskirala niska tražnja.

U poređenju sa fiksnim tarifnim mehanizmima, osnovni nedostatak *spot pricing*-a sa aspekta *cloud* provajdera je složenost i dodatni izazovi u pogledu razvoja adekvatnog tržišno orijentisanog mehanizma. Ukoliko se pretpostavi da je *spot pricing* definisan kao aukcija, bez obzira na to da li je cilj maksimizacija prihoda ili alokacija resursa, odgovarajući problem optimizacije bi bio NP-težak ili NP-kompletan. Pored toga, implementacija takvog tržišno orijentisanog mehanizma bi zahtevala dodatna ulaganja i troškove. Netransparentne ponude u aukciji i međusobno dogovaranje mogu uzrokovati ciklični efekat i dovesti do toga da takva tržišno orijentisana strategija postane neprofitabilna za *cloud* provajdere. Zapravo, *spot pricing* se smatra suviše jednostavnim mehanizmom da bi garantovao istinite ponude u aukciji i fer alokaciju resursa. Takođe, primena *spot pricing*-a može stvoriti cenovnu diskriminaciju i ostvariti niže prihode od fiksnih mehanizama [9]. Postoje takođe i spekulacije o ispravnosti politike Amazona, kao najvećeg *cloud* provajdera, o nenaplaćivanju korisnicima prekinutog parcijalnog sata korišćenja *cloud* resursa ukoliko je do prekida došlo od strane Amazona. U najgorem slučaju, neplaćeni resursi povećavaju do 30% od ukupnog vremena rada *spot* instanci, što predstavlja dodatni trošak i smanjuje profit [9]. Ipak, obzirom da je ovo Amazonova sopstvena tarifna politika, ta ograničenja se ne smatraju nedostacima *spot pricing*-a kao takvog.

Iako *spot pricing* u opštem slučaju može neočekivano povećati cene iznad *on-demand* cena, uvreženo je mišljenje da su *spot* instance jeftinije. Kvantitativne analize cena Amazona pokazuju da korisnici mogu ostvariti i do upola veće uštede zamenom *on-demand* instanci *spot* instancama. Empirijske studije pokazuju da uz adekvatne aukcijske ponude, ukupni troškovi iznajmljivanja *spot* instanci mogu biti od 13% do 36% od vrednosti troškova od iznajmljivanja ekvivalentnih *on-demand* instanci. *Spot pricing* tarifni mehanizam može povećati iskustveni kvalitet korisnika pri obezbeđivanju *cloud*

resursa. Najpre, ako postoji nedostatak *on-demand* resursa, *cloud* korisnici mogu skratiti vreme čekanja na opslugu prelaskom na alternativne *spot* instance. *Spot* instance su pogodne i za obezbeđivanje korisničkih zahteva manjeg obima. Takođe, ako se u okviru *spot pricing*-a primeni mehanizam kombinatornih aukcija, manipulisanje različitim tipovima *cloud* resursa postaje značajno efikasnije i pogodnije za korisnike sa složenim zahtevima. Tržišno orijentisano dinamičko obezbeđivanje resursa ostvaruje bolji *social welfare* i fer interakcije između *cloud* provajdera i korisnika.

Najveći izazov kada se radi o obezbeđivanju *cloud* resursa je upravo dostupnost istih. U skladu sa principima *spot pricing*-a, *cloud* provajder ima kontrolu terminiranja *spot* servisa kroz tarifna podešavanja. Drugim rečima, *spot* instance mogu postati nedostupne u bilo kom trenutku bez znanja korisnika usled varijacija u ponudi i tražnji. Pokazano je da se na osnovu istorije *spot* cena Amazona može doći do otkaza obezbeđivanja *cloud* resursa čak i kada je ponuda datog korisnika ista kao cena regularnih instanci [11, 17-18]. Samim tim, teško je analizirati dostupnost *spot* instanci obzirom na moguća velika kašnjenja. Sa aspekta krajnjeg korisnika, iskustveni kvalitet se značajno može ugroziti usled neizvesnih performansi što neminovno dovodi do negativnih efekata na sve prethodno pomenute prednosti primene *spot pricing*-a. Neispunjavanje garantovanog nivoa kvaliteta servisa dovodi do plaćanja penala. Pored toga, ukoliko nisu adekvatno definisane, niže cene *spot* instanci sa čestim otkazima mogu biti zapravo skuplje u odnosu na *on-demand* instance usled značajno viših troškova. Iz tih razloga *spot pricing* se smatra neadekvatnim za manje fleksibilne zahteve korisnika za obezbeđivanje servisa, a naročito u slučaju zahteva sa dužim trajanjem opsluge. Nerazumevanje od strane korisnika je još jedan nedostatak *spot pricing*-a. Korisnici se često suočavaju sa dilemama pri donošenju odluka u pogledu visine ponude u aukciji. Sa jedne strane, izuzetno niska ponuda može rezultirati u dužem vremenu opsluge bez značajno nižih troškova. S druge strane, izuzetno visoka ponuda može značajno uvećati troškove bez značajnog smanjenja vremena opsluge.

5. Zaključak

Intenzivan razvoj *cloud computing*-a i balansiranje ponude i tražnje za *cloud* resurse na osnovu dinamičkih mehanizama tarifiranja sve više dobija na značaju. Dinamičko tarifiranje predstavlja obećavajući mehanizam za unapređenje poslovnih modela *cloud* provajdera. Ovakav tarifni mehanizam omogućava optimizaciju iskorišćenja resursa kao i optimizaciju troškova javnih *cloud* provajdera. *Cloud* provajderi poseduju velike kapacitete koji većinom ostaju neiskorišćeni čime provajderi snose gubitke. Iz tog razloga, teži se podsticaju korisnika da iznajmljuju te kapacitete i povećaju tražnju. Ovo se smatra prvim korakom ka potpunom kreiranju tržišta za računarske resurse, *spot* tržišta. *Spot* instance su virtualne mašine dostupne korisnicima po nižim cenama u odnosu na standardne *on-demand* instance. Sam mehanizam se zasniva na aukcijama, gde korisnici kreiraju svoje ponude za date *spot* instance. Ove instance su aktivne sve dok je trenutna cena niža od maksimalne ponude korisnika koja predstavlja iznos koji je korisnik spreman da plati za čas korišćenja date instance. Uprkos brojnim prednostima koje omogućava *spot pricing* u ekonomskom pogledu, zbog promenljivih uslova na *spot* tržištu i izazova u vidu raspoloživosti resursa, terminiranja instanci, otkaza i nejasnoća koje prate ovaj tarifni mehanizam sa aspekta korisnika, ovaj mehanizam još uvek nije doživeo široku primenu.

Zahvalnica

Ovaj rad je deo istraživanja u okviru projekta TR32025 Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije.

Literatura

- [1] J. Lee, B. K. Szymanski, "A Novel Auction Mechanism for Selling Time-sensitive e-services", *Proceedings of 7th IEEE International Conference on E-Commerce Technology, (CEC'05)*, Hong Kong, 2005, pp. 75–82.
- [2] A. N. Toosi, R. N. Calheiros, R. K. Thulasiram, R. Buyya, "Resource Provisioning Policies to Increase IaaS Provider's Profit in a Federated Cloud Environment", *Proceedings of the 13th IEEE International Conference on High Performance Computing and Communications (HPCC'11)*, Banff, Canada, 2011, pp. 279-287.
- [3] M. Marcias, J. Guitart, "A Genetic Model for Pricing in Cloud Computing Markets", *Proceedings of the 2011 ACM Symposium on Applied Computing*, Taiwan, 2011, pp. 113-118.
- [4] A. N. Toosi, K. Vanmechelen, F. Khodadadi, R. Buyya, "An Auction Mechanism for Cloud Spot Markets", *ACM Transactions on Autonomous and Adaptive Systems*, vol. 11, no. 1, February 2016, pp. 25-57.
- [5] S. Kansal, G. Singh, H. Kumar, S. Kaushal, "Pricing Models in Cloud Computing", *Proceedings of the First International Conference on Information and Communication Technology for Competitive Strategies*, Udaipur, India, 2014.
- [6] ITU-T Recommendation Y.3500, "Information technology – Cloud computing – Overview and vocabulary", 2014.
- [7] V. K. Singh, K. Dutta, "Dynamic Price Prediction for Amazon Spot Instances", *Proceedings of the 48th Hawaii International Conference on System Sciences*, Hawaii, 2015.
- [8] Z. Li, H. Zhang, L. O'Brien, S. Jiang, Y. Zhou, M. Kihl, R. Ranjan, "Spot Pricing in the Cloud ecosystem: A Comparative Investigation", *The Journal of Systems and Software*, vol. 114, 2016, pp. 1-19.
- [9] V. Abhishek, I. A. Kash, P. Key, "Fixed and Market Pricing for Cloud Services", *Proceedings of the 7th Workshop on the Economics of Networks, Systems, and Computation (NetEcon2012)*, Orlando, USA, 2012, pp. 157-162.
- [10] M. Stokely, J. Winget, E. Keyes, C. Grimes, B. Yolken, "Using a Market Economy to Provision Compute Resources across Planet-wide Clusters", *Proceedings of IEEE International Symposium on Parallel Distributed Processing (IPDPS'09)*, Rome, Italy, 2009.
- [11] B. Javadi, R. K. Thulasiram, R. Buyya, "Characterizing Spot Price Dynamics in Public Cloud environments", *Future Generation Computer Systems*, vol. 29, no. 4, June 2013, pp. 988–999.
- [12] L. M. Leslie, Y. C. Lee, P. Lu, A. Y. Zomaya, "Exploiting Performance and Cost Diversity in the Cloud", *Proceedings of the 6th IEEE International Conference Cloud Computing (CLOUD2013)*, Santa Clara, USA, June 2013, pp. 1-19.
- [13] S. Zaman, D. Grosu, "A Combinatorial Auction-based Mechanism for Dynamic VM Provisioning and Allocation in Clouds", *IEEE Transactions on Cloud Computing*, vol. 1, no. 2, 2013, pp. 129–141.

- [14] V. Kantere, D. Dash, G. Francois, S. Kyriakopoulou, A. Ailamaki, "Optimal Service Pricing for a Cloud Cache", *IEEE Transactions on Knowledge and Data Engineering*, vol. 23, no. 9, September 2011, pp. 1345–1358.
- [15] M. Karakus, Z. Li, W. Cai, T. N. B. Duong, "OMTiR: Open Market for Trading Idle Cloud Resources", *Proceedings of the 6th IEEE International Conference on Cloud Computing Technology and Science (CloudCom2014)*, Singapore, 2014, pp. 719–722.
- [16] Y. Zhang, B. Li, Z. Huang, J. Wang, J. Zhu, H. Peng, "Strategy-proof Auction Mechanism with Group Price for Virtual Machine Allocation in Clouds", *Proceedings of the 2nd International Conference on Advanced Cloud and BigData (CBD2014)*, China, 2014, pp. 60–68.
- [17] S. Chaisiri, R. Kaewpuang, B. Lee, D. Niyato, "Cost Minimization for Provisioning Virtual Servers in Amazon Elastic Compute Cloud", *Proceedings of the 19th Annual IEEE International Symposium on Modelling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS2011)*, Singapore, July 2011, pp. 85–95.
- [18] S. H. Lim, G. S. Thakur, J. L. Horey, "Analyzing Reliability of Virtual Machine Instances with Dynamic Pricing in the Public Cloud", *Proceedings of the 28th IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW2014)*, Phoenix, USA, May 2014, pp. 885–893.

Abstract: *Comprehensive development of cloud computing leads to necessity of appropriate pricing plans' improvement. Fixed pricing plans, characterized by fixed prices in time unit, are widespread applied. In order to seize all cloud computing advantages, dynamic pricing plans depicting relations among supply and demand, are needed. This paper analyzes the possibility of application of spot pricing plan. Basic characteristics of spot pricing and auctions as a realization of spot pricing are presented. Advantages and disadvantages of this pricing plan are analyzed from cloud provider's perspective and cloud customer's perspective, as well.*

Keywords: *spot pricing, cloud computing, cloud providers, cloud services, cloud customers*

POSSIBILITY OF APPLICATION OF SPOT PRICING MECHANISM IN CLOUD ENVIRONMENT

Branka Mikavica, Vesna Radonjić Đogatović, Aleksandra Kostić-Ljubisavljević