

**BIZNIS ANALIZA MODELA UVODJENJA CLOUD SISTEMA U PREDUZEĆE
BUSINESS ANALYSIS OF A TRANSITION MODEL TO A CLOUD SOLUTION WITHIN AN ENTERPRISE**

Vladimir Đurica

REZIME: U okviru rapidno evolutivnog i konkurentnog poslovanja, kompanije i njihov menadžment, prinuđeni su na brzu adaptaciju tržišnih promena i usredsređenje na stvaranje nove vrednosti te ostvarenja maksimizacije povraćaja ulaganja u znanje i tehnologiju. Da bi kompanije postigle uspeh u novom okruženju, moraju restrukturirati postojeće organizacione strukture, redefinisati svoje strategije, poslovne procese, i izgraditi tehnološku infrastrukturu neophodnu radi efikasnijeg poslovanja i većeg povraćaja investicija ili možda unajmiti potrebne resurse. Iako koncept najma resursa nije nov, sa pojavom cloud sistema ovaj koncept se prenosi i na sferu najma računarskih usluga. Koncept ide do tog nivoa apstrakcije da neko ko nudi softver kao uslugu može to da izgradi korišćenjem određene platforme ili infrastrukture na najam. Tako, dolazimo do cloud sistema na cloud sistemu. Dalje, postoje javni i privatni cloud sistemi. Ukoliko nemamo dovoljno resursa u svom privatnom, možemo da se proširimo na javni za onoliko vremena i resursa koliko nam je to potrebno. Rad se bavi pristupom nalaženju optimalnog cloud rešenja, kao i razmatranjem poslovnih izbora hipervizora, kao osnove za izgradnju cloud sistema.

KLJUČNE REČI: Poslovni razlozi za prelazak na Cloud, cloud, virtualizacija, SPI model, SaaS, PaaS, IaaS

ABSTRACT: With rapidly evolving and competitive market situation, companies and their management are forced to a fast adaptation to the market changes and hence to focus on discovering ways of creating new value and maximizing the return on investments in knowledge and technology. If the companies could succeed in new surrounding they must change current organizational structures, redefine their strategies, business processes, build technology infrastructure or perhaps contract as a service, necessary in order to support efficiently business processes and provide return on investments. Although concept of renting resources is not new, with emerging of cloud systems this concept is going to a sphere of computing services. This concept goes to the extent of someone offering software as a service can use platform or infrastructure as an underlying service. This is a cloud on top of a cloud system. Also, it is possible to combine public and private cloud system. If we have insufficient resources in the private cloud, we can expand to public cloud for as much resources are required. This paper is focused on finding optimal cloud solution on a business level.

KEY WORDS: Business reasons for switching to Cloud solution, cloud, virtualization, SPI model, SaaS, PaaS, IaaS

UVOD

Postavlja se pitanje da li je ekonomičnije preći sa datacentra na cloud hostovano rešenje, ili pak zadržati se na postojećem rešenju i kako doći do saznanja u kojoj tački je to isplativo? Takođe, razmatraćemo kasnije i aspekte koji se odnose na evaluaciju izgradnje sopstvene infrastrukture naspram najma resursa. Konačno, treći deo razmatranja se odnosi na postavljanje poslovnih kriterijuma za izbor hipervizora.

Tabela 1 sadrži ažuriranovu Grejevu tabelu iz 2003. Godine, na podatke iz 2008., omogućavajući nam da pratimo stopu promene cloud tehnologija tokom ovih pet godina. Primetimo da su cene širokoprojasnog pristupa Internetu ponajmanje opale, ispod faktora 3. Troškovi računarske obrade opali su najviše, dok se sposobnost iskorišćenja dodatne računarske snage zasniva na pretpostavci da program može da iskoristi sva procesorska jezgra u okviru računarskog sistema. Ova pretpostavka je pre tačna za servisno računarstvo, sa puno virtualnih računara koji opslužuju hiljade ili milione korisnika, nego za programe koji se izvršavaju unutar datacentra jedne kompanije.

Da bi olakšao izračunavanje, Grej je izračunao šta je 1 USD mogao da kupi u 2003 [1]. Godini. Tabela 1 pokazuje njegove brojke naspram 2008. godine. i upoređuje sa EC2/S3 troškovima. Na prvi pogled, izgleda da će dolar više dati ukoliko se upotrebi za kupovinu hardvera u 2008. godini nego da se plati za korišćenje istog hardvera. S druge strane, ova jednostavna analiza prenebreže nekoliko bitnih faktora [2].

Posebno plaćanje po osnovu resursa. Najveći broj aplikacije nejednako koriste resurse za obradu, skladištenje i mrežni saobraćaj; neke su vezane za procesore, neke za mrežne resurse, i tako dalje, i mogu zasititi te resurse dok druge ne koriste u dovoljnoj meri. Model naplate na osnovu korišćenja svojstven cloud servisima omogućava posebno plaćanje po osnovi upotrebe datog resursa. Na ovaj način se smanjuje količina neiskorišćenosti resursa koji su na raspolaganju. Dok tačne uštede zavise od aplikacije, pretpostavimo da se za određen slučaj procesorska snaga koristi u obimu od oko 50%, dok su mrežni resursi u potpunosti iskorišćeni; u tom slučaju imamo za scenario datacentra plaćanje dvostruko više u odnosu na realno iskorišćenje procesorskih resursa. Tako, umesto da kažemo da 2.65 USD košta da se iznajmi procesor vredan 1 USD, bilo bi tačnije reći da košta 2.65 USD da se iznajmi procesor vredan 2 USD. Kao dodatna napomena, AWS cene za mrežni saobraćaj su zapravo konkurentnije od cene koju bi kompanija srednje veličine platila za isti saobraćaj.

	WAN propusni opseg mesečno	CPU sati (sva jezgra)	Skladište
Resurs u 2003	1 Mbps WAN link	2 GHz CPU/2 GB DRAM	200 GB disk, 50 Mb/transfer
Cena u 2003	100 USD	2000 USD	200 USD

1 USD kupuje u 2003	1 GB	8 CPU sati	1GB
Resurs u 2008	100 Mbps WAN link	2 GHz/2 socketa / 4 jezgra / 4 GB DRAM	1 TB disk, 115 MB/s transfer
Cena u 2008	3600 USD	1000 USD	100 USD
1 USD kupuje 2008	2.7 GB	128 CPU sati	10 GB
Cena/performanse	2.7x	16x	10x
Amazon instanca vrednosti 1\$ u 2008	\$0.27–\$0.40	\$2.56	\$1.20–\$1.50
Resurs u 2013	200 Mbps WAN link	2 GHz/8 jezgara/16 niti/20 GB DRAM	2 TB disk, 6GB/stransfer
Cena u 2013	360 USD	708 USD	72USD
1 USD kupuje 2013	43 GB	1024 CPU sati	22 GB
Cena/performanse	16x	8x	2.2x
Amazon instanca vrednosti 1\$ u 2013	\$0.04 - \$0.096	\$1.024	\$0.09 - \$0.11

Tabela 1 – Cena računarskih resursa u 2003, 2008 i 2013

Na Grejeve podatke iz 2003 dodati su podaci iz 2008 a zatim iz 2013, normalizovani u odnosu na kupovnu moć 1 USD 2003 naspram 2008, i upoređena je sa cenom najma 1 USD vrednih resursa na AWS za 2008. godinu i 2013.

Struja, hlađenje i fizički troškovi fabrike. Cena struje, hlađenja i amortizacije zgrade su isključeni za sada iz ovog pojednostavljenog modela. Hamilton procenjuje da cena procesora, stordiza i mrežnog propusnog opsega se okvirno duplira ukoliko se uračuna amortizacija zgrade[3]. Koristi procenu, kupovinom 128 sati procesorskog vremena u 2008 stvarno košta 2 USD pre nego 1 USD, u poređenju sa 2.56 na EC2. Slično, 10 GB prostora na disku košta 2 USD pre nego 1 USD u poređenju sa 1.20 – 1.50 USD na mesečnom nivou na S3 Amazonovom servisu. Na kraju, S3 zapravo replicira podatke bar 3 puta radi trajnosti i performansi, obezbeđuje trajnost, i repliciraće podatke dalje ukoliko postoji visoka tražnja za podacima. To zapravo znači da je cena 6.0 USD kada se kupuje naspram 1.20 do 1.50 USD mesečnog najma na S3.

Operativni troškovi (održavanje). Danas, troškovi upravljanja računarskim resursima su relativno niski. Ponovno pokretanje servera je jednostavno te i minimalno obučeno osoblje može da vrši zamenu komponenti u serverskim rek ormarima. S jedne strane, pošto servisno računarstvo koristi virtualne mašine umesto fizičkih, sa tačke gledišta cloud korisnika, ovakvi zadaci mogu biti povereni cloud dobavljaču usluga. S druge strane, zavisno od nivou virtualizacije, mnogi troškovi upravljanja softverom mogu ostati – nadogradnje, primena softverskih zakrpa, itd.

Vraćanje na temu “upravljanje naspram neupravljanom” infrastrukturom, veruje se da će cene biti niže u okvirima upravljanje infrastrukture (Microsoft Azure, Goolge AppEngine, Force.com) nego u domenu hardverskog nivoa isporuke infra-

strukture poput Amazon EC2 servisa. Ipak, izgleda veoma teško kvantifikovati dobrobiti na takav način da se stvori preovlađujuće mišljenje o ovim stvarima.

PRIMER 1: Migracija u cloud. Pretpostavimo da biološka laboratorija generiše 500 GB novih podataka prilikom svakog eksperimenta. Računaru EC2 intanciranom je potrebno 2 sata po GB da obradi te podatke. Laboratorija ima ekvivalent 20 instanci lokalno, tako da vreme da se evaluira eksperiment je 500 x 2 / 20 ili 50 sati. Mogli bi da obrade sve podatke za sat vremena na 1000 instanci AWS EC2 clouda. Cena ovakvog eksperimenta bila bi 1000 x 0.10 USD ili 100 USD u obradi podataka i još 500 x 0.10 USD = 100 USD za prenos podataka preko mreže. Dalje, mereno je vreme prenosa podataka od laboratorije do AWS-a pri brzini od 20 Mbita/sec (Garfinkel, 2007). Vreme prenosa je bilo 500 GB x 1000 MB/GB x 8 bita/byte) / 20 Mbita/sec = 4,000,000 / 20 = 200,000 sekundi ili više od 55 sati. Kada se to uporedi sa 50 sati lokalne obrade podataka, laboratorija je zaključila da im se ne isplati prelazak na cloud rešenje [4].

Povezan problem je softverska složenost i cena delimične ili pune migracije podataka sa tradicionalnog sistema na cloud. Pored toga što je migracije zadatak koji se obavlja jednom, može predstavljati velik izazov i trud. Ovaj zadatak predstavlja nove poslovne prilike za kompanije koje obezbeđuju integraciju podataka između javnog i privatnog cloud sistema.

EKONOMIJA CLOUD RAČUNARSTVA

Nekoliko napomena u vezi sa ekonomskim modelima cloud računarstva:

- U okviru odlučivanja da li se držanje servisa u cloudu isplati dugoročno gledano, čini se da fina granularnost ekonomskog modela u okviru cloud ponude čini ovo odlučivanje fluidnijim. Tako, na primer, osobina elastičnosti cloud sistema omogućava da se prelazak izvede uz prenos rizika.
- Takođe, iako cena hardverskih resursa ima tendenciju stalnog pada, to se takođe dešava u nepravilnom rasporedu; na primer, cena obrade i cena skladišta podataka brže padaju od cene WAN konekcija. Cloud računarstvo može da prati ove promene – i potencijalno ih prosledi do klijenta – efikasnije u odnosu na samostalni centar za obradu podataka koji bi taj isti klijent napravio, rezultujući bližoj ceni stvarnih troškova upotrebe tih resursa.
- Prilikom odlučivanja za pomeranje servisa u cloud, potrebno je dodatno utvrditi očekivan prosečni skok korišćenja resursa, posebno ako aplikacija može imati veoma promenljive skokove u zahtevima; praktični limiti u realnom svetu se nalaze na nivou instalirane opreme; različiti operacioni troškovi koji variraju u zavise od tipa cloud okruženja koje se razmatra.

Elastičnost: Pomeranje rizika

Iako se ekonomska privlačnost cloud računarstva često opisuje kao pretvaranje kapitalnih u operativne troškove

(CapEx to OpEx, eng. *Capital expenses to operative expenses*), čini se da je plati-na-osnovu-korišćenja dosta precizniji način da se opiše ekonomska dobrobit za zakupca cloud sistema. Ovako zakupljeni sati mogu biti neuniformno korišćeni tokom vremena a da se i dalje plaća po osnovu korišćenja (npr. koristi se 100 server sati danas a sutradan ništa, i da se i dalje plati samo korišćeno serversko vreme); u zajednici koja se bavi računarskim mrežama ovakav način prodaje je već poznat pod nazivom naplata po osnovu korišćenja. Dodatno, nepostojanjem kapitalnih izdvajanja za ovu namenu, kapital se može usmeriti na sam poslovni poduhvat.

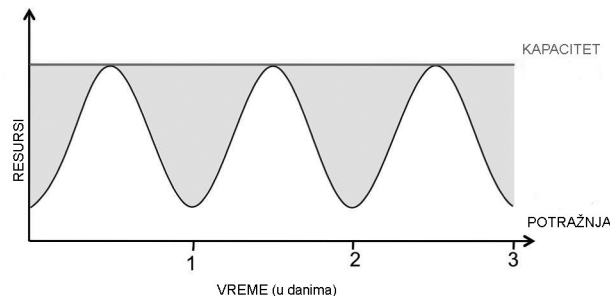
Tako, iako na primer, Amazonov sistem naplate na kraju može biti skuplji od kupovine i upotrebe uporedivog servera za isti period, cenovna razlika prevaže u korist cloud koncepta koji nudi ekonomske prednosti iz kategorije elastičnosti sistema i transfera rizika, naročito rizika od prevelike ili nedovoljne alokacije resursa.

Počecemo sa elastičnošću. Ključno zapažanje je da osobina cloud računarskog sistema da se dodaju ili uklanjaju resursi uz finu granularnost (po jedan server u slučaju EC2) i sa vremenom uvođenja merenom u minutima, u poređenju sa nekoliko nedelja, omogućava poklapanje alociranih resursa sa potrebom za poslom koji ti serveri treba da obavie. U praktičnoj upotrebi, iskorišćenje servera u datacentrima je u rasponu od 5% do 20% [5]. Ovo, iako ovakav stepen iskorišćenost deluje nisko, proističe iz uočavanja da u odnosu na standardnu upotrebu, prilikom povećanih zahteva potrebe skaču sa faktorom množenja u rasponu od 2 do 10. Pojedini korisnici namerno obezbede manje resursa nego što je potrebno u periodu maksimalne potražnje i dozvoljavaju resursima da budu neupotrebljeni u vreme kada za njima nema potražnje. Što su varijacije izraženije to ima više neupotrebljenih resursa. Jednostavan primer objašnjava kako elastičnost jednog ovakvog sistema omogućava više nego dovoljno kompenzaciju više cene po server/satu sistema naplate po korišćenju.

PRIMER 2: Elastičnost. Pretpostavimo da imamo servis koji ima predvidivu dnevnu tražnju gde maksimum zahteva 500 servera a samo 100 tokom noći, kao što je prikazano u Slici 2(a). Dok god je prosečno iskorišćenje tokom dana 300 servera, stvarno iskorišćenje $300 \times 24 = 7200$ server sati; međutim, pošto moramo obezbediti 500 server sati u periodima maksimalnog iskorišćenja, plaćamo zapravo $500 \times 24 = 12\,000$ server sati, faktor 1.7 više nego što nam je potrebno. Zaključujemo da dok god god je cena sistema naplate po korišćenju tokom perioda od 3 godine manja od 1.7 puta cena koštanja servera, možemo ostvariti uštede korišćenjem servisnog računarstva [4].

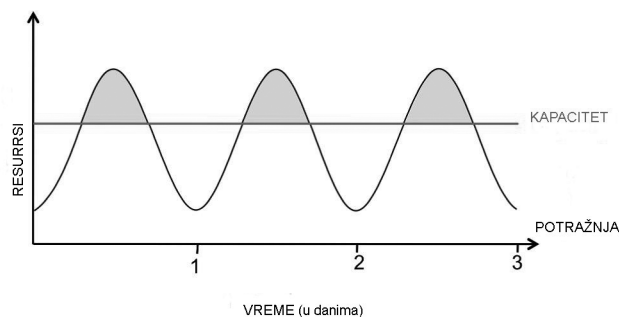
Prethodno naveden primer podcenjuje dobrobit elastičnosti jer kao dodatak na jednostavne dnevne obrasce upotrebe računarskih resursa, mnogi netrivialni resursi takođe imaju sezonske ili druge periodične varijacije u potrebi za resursima (npr. e-commerce sezonski skokovi i usluge deljenja slika nakon praznika i perioda godišnjih odmora) kao i neka neočekivana iskakanja zbog eksternih događaja poput vesti. Tradicionalno, obzirom da dobavljanje nove opreme može trajati nedeljama, jedini način da se ovakva potreba reši je

da se resursi obezbede unapred. Videli smo da, čak i kada dobavljači usluga dobro predvide ovakve skokove kapaciteti svakako ostaju neiskorišćeni, i ukoliko se preceni greškom količina potrebnih resursa, tim gore, Slika 1.

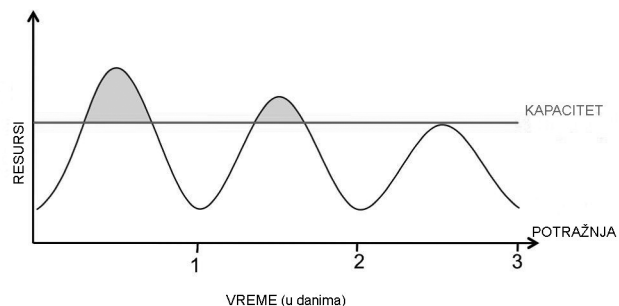


Slika 1 – Obezbeđivanje resursa u periodu povišene tražnje

Druga mogućnost, je svakako, da se podceni tražnja Slika 2, što bi moglo dovesti do odbijanja usluge za korisnike koji su preko kapaciteta sistema. Dok su finansijski efekti prebacivanja preko kapaciteta vrlo merljivi, efekte nedovoljne alokacije je vrlo teško meriti, ipak vrlo ozbiljno: ne samo da je prinos na odbijenog korisnika nula, već je vrlo moguće da je takav korisnik trajno odbijen. Slika 3 ima za cilj da predstavi ovakvo ponašanje: korisnici će napustiti sistem servisa koji nema dovoljno alociranih resursa u datom trenutku, do momenta do kog se ne smanji tražnja i resursi postanu ponovo slobodni. U tom momentu će se korisnici početi vraćati sistemu, međutim u manjem broju.



Slika 2 – Nedovoljno obezbeđenih resursa 1



Slika 3 – Nedovoljno obezbeđenih resursa 2

PRIMER 3: Transfer rizika. Pretpostavimo da je 10% korisnika koji imaju iskustvo malopre opisane loše usluge trajno izgubljeno, a da bi ti isti korisnici da su imali uslugu sve vreme ostali korisnici sistema – Sistem je inicijalno napravljen da

pruža usluge za 400,000 korisnika (1000 korisnika po serveru X 400 servera), međutim zbog pozitivnog novinskog članka tražnja poraste na 500,000 u prvom satu. Od 100,000 korisnika koji su odbijeni u smislu usluge (neusluženi), pretpostavka je da će njih 10,000 biti trajno izgubljeni kao klijenti, ostavljajući aktivnu korisničku bazu od 390,000. U narednom satu će sajt posetiti 250,000 novih korisnika ali sajt je i dalje za 240,000 korisnika preko svojih projektovanih kapaciteta. Ova situacija rezultira u dodatnih 24,000 odbijanja, ostavljajući bazu od 376,000 stalnih korisnika. Ukoliko se ovaj trend nastavi, posle 500,000 ili 19 sati, broj novih korisnika će se približiti nuli i sajt će biti u okviru kapaciteta. Jasno je da će ovako projektovan sistem sakupiti manje od 400,000 korisnika vrednih stabilnog prihoda tokom ovih 19 sati, međutim, da još jednom ilustrujemo argument nedovoljne alokacije resursa – da ne govorimo o lošoj reputaciji od strane nezadovoljnih korisnika.

Postavlja se pitanje da li se ovakvi scenariji zaista dešavaju u praksi? Kada je Animoto [6] učinio svoj servis dostupan putem Fejsbuka, desio se porast tražnje od 50 servera na 3500 servera za samo tri dana. Čak i da je prosečno iskorišćenje pojedinačnog servera bilo relativno malo, niko nije mogao predvideti dupliranje resursa na svakih 12 sati tokom tri dana.

Nakon što je smanjila tražnja, saobraćaj se sveo na znatno niži nivo. Tako, u ovom primeru, skaliranje sistema na gore i potreba za elastičnošću nije bila samo optimizacija troškova već i operativna potreba. Skaliranje na dole je omogućilo da se iskorišćenje resursa na čekanju tačnije podudara sa potrebama. Elastičnost je od vrednosti kako postojećim firmama, tako i novopokrenutim biznisima. Na primer, firma "Target", američki drugi po veličini gigant maloprodaje, koristi AWS za Target.com web sajt. Dok su drugi dobavljači imali znatne probleme sa performansama, povremenu nedostupnost na "crni petak" (28. Novembar), Target.com i Amazon.com sajtovi su bili sporiji za samo 50%. Koristeći isti princip elastičnosti, Salesforce.com hostuje klijente od dva do 40,000 korisnika.

Čak i manje dramatični slučajevi sasvim su dovoljni radi ilustracije ključnih dobrobiti od cloud računarstva: rizik od pogešne procene zahteva ka sistemu se pomera od dobavljača usluge ka vlasniku cloud sistema. Cloud vlasnik može da naplati dodatatak za preuzimanje ovog rizika koji bi zapravo reflektovao višu cenu upotrebe server sata u poređenju sa trogodišnjom cenom kupovine istog sistema. Postoji prosta jednačina koja uopštava sve prethodno pomenute slučajeve. Pretpostavlja se da cloud dobavljač usluga uposli sistem koji se plaća po osnovu korišćenja, u okviru kojih korisnik plaća proporcionalno vreme i količinu resursa koje koristi. Dok izvesni [7], [8], [9] smatraju da je potreban sofisticiraniji model plaćanja za usluge infrastrukture, izgleda da će princip postavljanja cene po osnovu korišćenja biti preovlađujući jer je jednostavniji i transparentniji, kao što je slučaj sa širokom primenom "pravih" uslužnih servisa poput struje, gasa, itd. Takođe, cena bi se mogla redukovati manipulacijom u okviru isporuke sadržaja [13]. Slično tome, pretpostavlja se da je klijentov profit direktno proporcionalan broju upotrebljenih sati. Ova pretpostavka je konzistentna sa online oglašavačkom politikom generisanja profita u kojoj je broj prikaza oglasa

direktno proporcionalan ukupnom vremenu poseta provedenom od strane krajnjih korisnika na servisu [4].

$$\text{KorisnikSat}_{\text{cloud}} \times (\text{prihod} - \text{Cena}_{\text{cloud}}) \geq \text{KorisnikSat}_{\text{datacenter}} \times (\text{prihod} - \text{Cena}_{\text{datacenter}} / \text{Iskorišćenje})$$

Leva strana jednačine množi neto prihod po korisnik satu (prihod realizovan po korisnik satu umanjeno za cenu plaćanja cloud računara po korisnik satu) dajući očekivani profit od korišćenja cloud sistema. Desna strana izvršava istu kalkulaciju samo za datacenter fiksnog kapaciteta množenjem prosečnog korišćenja, uključujući i periode bez maksimalne tražnje. Koja god strana je veća, predstavlja potencijalni profit. Očigledno, ako je iskorišćenje = 1:0 (oprema datacentra ima 100% iskorišćenja), u kom slučaju obe strane jednačine izgledaju isto. Ipak, teorija redova čekanja nam govori da kako se približavamo stanju 1:0, vreme odziva sistema se približava beskonačnosti. U praksi, korisni kapacitet datacentra (bez kompromitovanja usluge) je obično između 0.6 i 0.8. Dok datacenter mora da dobavi više resursa radi obezbeđivanja usluge u momentima maksimalne tražnje, cloud dobavljač jednostavno može da uključi ovu vrstu rizika u cenu najma resursa. Ovo praktično znači da ukoliko iznajmite Internet link od 100 Mbita/sekundi, moći ćete da koristite svega 60 do 80 Mbita/sekundi u praksi.

Jednačina jasno pokazuje da je zajednički element u svim primerima sposobnost kontrolisanja cene po korisnik satu operativnog servisa. U primeru 2, cena po korisnik satu bez elastičnosti je visoka zbog resursa koji stoje neiskorišćeni – viša cena dok je isti broj korisnik sati. Ista stvar se dešava ukoliko se preceni zahtev, te alociraju resursi koji kasnije ostaju neiskorišćeni. U drugom primeru, cena po korisnik satu se povećava kao rezultat podcene maksimalnog skoka zahteva koji rezultira odbijanjem korisnika: Obzirom da se neki procenat korisnika trajno gubi, fiksni troškovi ostaju isti ali se sada amortizuju manjim brojem korisnik sati. Ovo ilustruje fundamentalna ograničenja sistema "kupovnog" modela u kontekstu bilo kakvog netrivialnog povećanja zahteva.

Na kraju, postoje dve dodatne koristi iz cloud sistema za korisnika iz mogućnosti da promeni svoju upotrebu resursa u satima, naspram godina. Prvo, neočekivano skaliranje na dole (oslobađanje trenutno neupotrebljenih resursa) – na primer, tokom poslovnog usporavanja, ili ironično tokom poboljšane softverske efikasnosti – standardno ovakve situacije nose finansijsku kaznu sa sobom. Uz trogodišnju amortizaciju, server od 2,100 USD, ukoliko se iz navedenih razloga izopšti iz sistema posle samo godinu predstavlja "kaznu" od 1,400 USD. Cloud računarstvo eliminiše ovu kaznu.

Drugo, tehnološki trendovi sugerišu da se tokom korisnog dela životnog ciklusa neke opreme, cena hardvera pada i novi hardver i softver postaju dostupni. Cloud provajderi, koji već uživaju u dobrobiti zahvaljujući velikom obimu nabavke, potencijalno mogu preneti neke od ovih ušteda na svoje klijente. Zaista, ozbiljni korisnici AWS sistema su imali prilike da osete pad cena za 20% u domenu sistema skladištenja podataka dok je u oblasti mreža pad bio 50% tokom proteklih 2.5 godine, uz dodatak devet novih servisa tokom perioda od

manje od jednu godinu. Ukoliko nove tehnologije ili cenovni planovi postanu raspoloživi dobavljaču cloud usluga, postojeće aplikacije i klijenti mogu potencijalno imati koristi od ove vrste inovacije momentalno, bez potrebe za kapitalnim ulaganjima u cilju uštede. Za manje od dve godine, Amazon Web Servisi su povećali broj različitih tipova servera ("instanci") sa jedne na pet, i za manje od jedne godine je dodato sedam novih infrastrukturnih servisa i dva nova načina za podršku.

Poslovni kriterijumi izbora hipervizora kao osnove za Cloud

KRITERIJUM SPOSOBNOSTI IZVRŠAVANJA

Kriterijum sposobnosti izvršavanja se bazira na evaluaciji kvaliteta i efikasnosti procesa, sistema, metoda ili procedura koje omogućavaju ICT dobavljačima usluga performanse da budu kompetitivni, efikasni i efektivni sa sveukupno pozitivnim uticajem na prihod zadržavanje i reputaciju. Na kraju, tehnološki dobavljači su evaluirani po osnovu sposobnosti da kapitalizuju svoju viziju [10, 11].

Kriterijum evaluacije	Težina
Proizvod / Usluga	Visoka
Opšta vitalnost (Poslovnih jedinica, Finansije, Strategija, Organizacija)	Visoka
Realizovanih prodaja / Politika cena	Visoka
Tržišna responzivnost i praćenje	Niska
Marketing	Visoka
Korisničko iskustvo	Standardno
Operacije	Nisko

Tabela 2 - Evaluacioni kriterijumi povodom sposobnosti realizacije

Sposobnost izvršavanja u domenu virtualizacije serverske infrastrukture nije vezano isključivo za funkcionalnosti proizvoda, već u velikoj meri za održavanje konstantno promenljivog poslovnog modela sa veoma dinamičnim trendom. Odlični proizvodi mogu propasti, ali i loši proizvodi mogu biti veoma uspešni, zavisno od efektivne proizvođačke realizacije.

Proizvod ili usluga: Ključni proizvodi i usluge u ponudi tehnološkog ponuđača koji zaokružuju određeno tržište. Ovo uključuje trenutne proizvodne i uslužne kapacitete, kvalitet, skup funkcionalnosti, veštine, itd., bilo da su u ponudi direktno ili kroz OEM i partnerske ugovore. Ključni faktori koji se evaluiraju uključuju: opseg operativnih sistema, skalabilnost i efikasnost, elastičnost, zrelost, upravljanje embedded resursima, upravljanje opcijama za efikasniju administraciju, mogućnost administriranja sistema kao celine, virtualizovanih ekosistema, administrativna skalabilnost i integracija se drugim dobavljačima usluge upravljanja i održavanja ovakvih sistema.

Opšta vitalnost (Poslovnih jedinica, Finansije, Strategija, Organizacija): Procena celokupne finansijske vitalnosti organizacije, finansijski i praktičan uspeh poslovnih jedinica i verovatnoća individualnih poslovnih jedinica kako bi se nastavilo sa investicijom u proizvod, nastavak nuđenja i podrške za proizvod i nastavak ponude tehnološki najsofisticiranijeg proizvoda u okviru tehnološkog proizvodnog portfolia kompanije.

Realizovanih prodaja / Politika cena: Mogućnosti tehnološkog dobavljača usluga u svim predprodajnim aktivnostima, i struktura koja ih podržava. Ovo uključuje upravljanje sklapanjem poslovima, cenovna politika i pregovaranje, predprodajna podrška i celokupna efektivnost prodajnih kanala.

Tržišna responzivnost i praćenje: Sposobnost odgovora, promene pravca, fleksibilnosti i konkurentnosti sa razvojem mogućnosti, konkurentnog delovanja, evolucije korisničkih zahteva i tržišne dinamičke promene. Ovaj kriterijum takođe razmatra dobavljačevu istoriju i responzivnost.

Marketing: Jasnost, kvalitet, kreativnost i efikasnost programa dizajniranih da isporuče poruku koju organizacija želi da komunicira i na taj način utiče na tržište, promocija brenda i posla, povećanje svesnosti vezane za proizvod, i ustanovljene pozitivne identifikacije sa proizvodom i/ili brendom od strane kupaca. Ovo je moguće postići kombinacijom publikovanja, promotivnih aktivnosti, liderskim nastupom, ličnom preporukom i proizvodnih aktivnosti.

Korisničko iskustvo: Odnosi, proizvodi i usluge/programi koji omogućuju klijentima da budu uspešni sa proizvodima koji se evaluiraju. Specifično, ovo uključuje načine na koji korisnici primaju tehničku i ostalu podršku. Ovo takođe može uključivati pomoćne alate, programe korisničke podrške (te i kvalitet), dostupnost korisničkih grupa, ugovore servisne podrške, itd.

Operacije: Sposobnost organizacije da ispuni svoje ciljeve i obaveze. Faktori uključuju kvalitet organizacione strukture poput veština, iskustava, programa, sistema i drugih elemenata potrebnih organizaciji kako bi efektivno i efikasno nastavila da funkcioniše.

Kriterijum evaluacije	Težina
Razumevanje tržišta	Visoka
Marketing strategija	Visoka
Prodajna strategija	Standardna
Ponuda (proizvodne) strategije	Standardna
Bizis Model	Standardna
Vertikalna / Industrijska strategija	Standardna
Inovacija	Standardna
Geografska strategija	Niska

Tabela 3 - Kriterijum evaluacije, kompletnost vizije

KRITERIJUM KOMPLETNOSTI VIZIJE

Kriterijumi kompletnosti vizije odnose se na sposobnosti tehnoloških dobavljača da ubedljivo artikuliraju logičke izjave povodom trenutnih i budućih tržišnih pravaca, inovacija, korisničkih potreba i kompetitivnih snaga. Na kraju, tehnološki dobavljači usluga se rangiraju po svom razumevanju kako se tržišne snage mogu iskoristiti da kreiraju nove mogućnosti za dobavljače.

U domenu serverske virtualizacije, razumevanje dobavljača tehnoloških usluga i njegova artikulacija strateških tehnoloških

kih putanja vezanih za virtualizaciju je naročito bitno. Ovo se odnosi na širenje osnova buduće infrastrukture i njena dalja moguća ekspanzija ka cloud računarstvu je od velikog značaja i to ga diferencira.

Razumevanje tržišta: Sposobnost tehnološkog dobavljača da razume potrebu svojih klijenata i da ih prevede proizvode i usluge. Dobavljači usluga koji pokazuju nivo vizije slušaju i shvataju šta klijenti žele i šta im treba i mogu da oblikuju ili unaprede te želje sa sopstvenim dodatkom vizije.

Marketing strategija: Jasnan i diferenciran skup poruka konzistentno komuniciran kroz organizaciju i eksteralizovan kroz web sajt, reklame, korisničke programe i izjave za pozicioniranje.

Podajna strategija: Strategija prodaje proizvoda koji koriste odgovarajuću mrežu direktne i indirektno prodaje, marketinga, usluga i komunikacionih posrednika koji šire okvire dohvatanja tržišta, veština, ekspertize, tehnologije, usluga i baze korisnika.

Ponuda (proizvodne) strategije: Pristup tehnološkog dobavljača razvoju proizvoda i isporuci koja uključuje strategiju diferencijacije, funkcionalnosti, metodologije i skupa funkcionalnosti, kako se mapiraju na trenutne i buduće zahteve.

Biznis model: Čvrstoća i logika fundamentalne logike poslovne ponude tehnološkog dobavljača.

Vertikalna / Industrijska strategija: Strategija tehnološkog dobavljača da upravlja resursima, veštinama i ponudama kako bi se odgovorilo na specifične potrebe ili pojedinačne tržišne segmente, uključujući vertikalne.

Inovacija: Direktni, povezani, komplementarni i sinergijski raspored resursa, ekspertiza ili kapitalna investicija, konsolidacija, odbrambena ili preventivna svrha.

Geografska strategija: Strategija tehnološkog dobavljača da upravlja resursima, veštinama i ponudama izvan "kuće" ili svoje matične geografije, bilo direktno ili preko partnera, kanala i predstavništava na odgovarajuć način u kontekstu geografije i tržišta.

ZAKLJUČAK

Napokon se ostvaruje dugo očekivana situacija da se računarski resursi mogu najmiti na osnovu njihove upotrebe. Princip elastičnosti se upoređuje sa poslovnom potrebom i služi direktno klijentima putem Interneta, kao što količina posla može da naraste (ili se smanji) znatno brže nego pre 20 godina. Nekada je bilo potrebno mnogo godina da bi posao narastao do nekoliko miliona klijenata – sada je to moguće mereno u mesecima.

Sa tačke gledišta cloud dobavljača, pravljenje veoma velikog centra za obradu podataka, cloud centra, na mestima gde je gradnja jeftina, korišćenjem jeftinih komponenti, storidža, i njihovim umrežavanjem, stvara se mogućnost prodavanja istih resursa baziranih na njihovom korišćenju ispod cene data-centra prosečne veličine, dok se profit multipleksira u okviru velikih grupa klijenata. Iz perspektive korisnika cloud sistema, interesantno je novoosnovanim, tzv. startup softverskim

firmama, da prave sopstvene datacentre kao što bi to startup biznis uradio za fabričku liniju. Pored startup biznisa, mnoge postojeće organizacije koriste osobinu cloud elastičnosti sistema redovno, uključujući primere kao što je Washington Post, Pixar, razni univerziteti.

Cloud sistemi su, iako ne revolucionarni po prepoznavanju tehnologija koje se u njima koriste, revolucionarni po poslovnom konceptu kojim se to isporučuje bilo krajnjem korisniku, bilo drugom dobavljaču rešenja kao međusloju.

Ovakav pristup će, nadamo se, omogućiti brži i kvalitetniji a pogotovo jednostavniji način na koji će se sistemi svih veličina graditi. Pri tome ne treba gubiti iz vida mnoge već predočene prednosti, nove mogućnosti koje postoje i one koje će tek biti uočene. Cloud je način da se optimalno izgradi sistem i drže pod kontrolom troškovi kako izgradnje, tako i upotrebe i održavanja.

REFERENCE

- [1] Gray, J. a. (2003). A conversation with Jim Gray. *ACM Queue* 1, , 8-17.
- [2] Gray, J. (2008). *Distributed Computing Economics*. Retrieved from http://portal.acm.org/ft_gateway.cfm?id=1394131&type=digital%20edition&coll=Portal&dl=GUIDE&CFID=
- [3] Hamilton, J. (2009). Cooperative Expendable Micro-Slice Servers (CEMS): Low Cost, Low Power Servers for Internet-Scale Services. *Conference on Innovative Data Systems Research*. CIDR.
- [4] Ambrust M., F. A. (2009). *Above the Clouds: A Berkeley View of Cloud Computing*. Retrieved from Berkeley College of Engineering: <http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf>
- [5] Rangan, K. (2008). The Cloud Wars: \$100 bilion at stake. *Tech Republic*. Merrill Lynch.
- [6] Bezos, Jeff. (2008, 04). Retrieved from Animoto: <http://blog.animoto.com/2008/04/21/amazon-ceo-jeff-bezos-on-animoto/>
- [7] Abramson, D. B. (2002). A computational economy for grid computing and its implementation in the Nimrod-G resource broker. *Future Generation Computer Systems* 18, 1061-1074.
- [8] Hosanagar, K. K. (2004). Optimal pricing of content delivery network (CDN) services. *The 37th Annual Hawaii International Conference on System Sciences* , 205-214.
- [9] Stuer, G. V. (2007). Acommodity market algorithm for pricing subsittutable Grid resources. *Future generation of computer systems* 23, 688-701.
- [10] Bittman, Thomas J. at all (2010). Magic Quadrant for x86 Server Virtualization Infrastructure. *Future generation of computer systems* 23, 688-701.
- [11] Magic Quadrant for x86 Server Virtualization Infrastructure (Gartner, 2013), <http://www.gartner.com/technology/reprints.do?id=1-1GJA88J&ct=130628&st=sb>
- [12] WebPass (2013), <http://www.webpass.net/residential>.
- [13] Deane, J. at all (2012), *Journal Information Technology and Management*, Volume 13 Issue 1, March 2012, Pages 1-15



Biografija: Autor je trenutno na doktorskim studijama Fakulteta Organizacionih Nauka Univerziteta u Beogradu gde je prethodno i magistrirao. Prethodnih 14 godina autor je radio u raznim domaćim i međunarodnim firmama radeći na implementaciji različitih ikt projekata.

Kontakt: djuricav@gmail.com, Daničareva 18, 11 000 Beograd, Republika Srbija.

Oblasti kojima se autor bavi: Softverski Definisane Mreže, Uvođenje cloud sistema, virtualizacija, mreže, operativni sistemi.