

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Andraž Glavan

Ekonomski in performančni vidik shranjevanja podatkov v oblaku

DIPLOMSKO DELO

VISOKOŠOLSKI STROKOVNI ŠTUDIJSKI PROGRAM PRVE
STOPNJE RAČUNALNIŠTVO IN INFORMATIKA

Ljubljana 2012

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Andraž Glavan

Ekonomski in performančni vidik shranjevanja podatkov v oblaku

DIPLOMSKO DELO

VISOKOŠOLSKI STROKOVNI ŠTUDIJSKI PROGRAM PRVE
STOPNJE RAČUNALNIŠTVO IN INFORMATIKA

Mentor: viš. pred. dr. Aljaž Zrnec

Ljubljana 2012



Št. naloge: 00234/2012

Datum: 04.04.2012

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Kandidat: **ANDRAŽ GLAVAN**

Naslov: **EKONOMSKI IN PERFORMANČNI VIDIKI SHRANJEVANJA
PODATKOV V OBLAKIH**

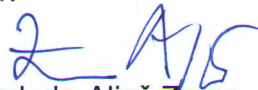
**ECONOMIC AND PERFORMANCE ASPECTS OF DATA STORAGE IN
THE CLOUDS**

Vrsta naloge: Diplomsko delo visokošolskega strokovnega študija prve stopnje

Tematika naloge:

Uporaba storitev za shranjevanje podatkov v oblakih omogoča podjetjem znižati skupne stroške lastništva strojne in programske opreme, ki je namenjena hrambi podatkov. V okviru poslovnega sistema lahko najdemo podatkovne zbirke, ki shranjujejo različne vrste podatkov: transakcijske podatke, dokumente in večpredstavnostne podatke. V okviru diplomske naloge se omejite na transakcijske podatke in dokumente v elektronski obliki ter preučite, katere vrste omenjenih podatkov bi bilo za potrebe poslovanja smiselno premestiti v oblak, pri čemer upoštevate predvsem performančne in ekonomske vidike.

Mentor:


viš. pred. dr. Aljaž Zrnec



Dekan:


prof. dr. Nikolaj Zimic

IZJAVA O AVTORSTVU

diplomskega dela

Spodaj podpisani/-a *Andraž Glavan*,

z vpisno številko *63050034*,

sem avtor diplomskega dela z naslovom:

Ekonomski in performančni vidik shranjevanja podatkov v oblaku

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal samostojno pod mentorstvom (naziv, ime in priimek)

viš. pred. dr. Aljaža Zrneca

- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identične s tiskano obliko diplomskega dela
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki »Dela FRI«.

V Ljubljani, dne *07.09.2012*

Podpis avtorja: _____.

ZAHVALA

Najprej bi se rad zahvalil svoji družini, za podporo v času študija in pisanja diplomske naloge. Zahvaljujem se tudi mentorju, viš. pred. dr. Aljažu Zrnecu za strokovne nasvete in pomoč, Ani Harb za lektoriranje in Žigu Glavanu za angleški prevod povzetka.

KAZALO

POVZETEK	1
ABSTRACT	3
1 UVOD	5
2 RAČUNALNIŠTVO V OBLAKU (ANGL. CLOUD COMPUTING)	7
2.1 INFRASTRUKTURA KOT STORITEV (IAAS).....	9
2.2 PLATFORMA KOT STORITEV (PAAS).....	9
2.3 PROGRAMSKA OPREMA KOT STORITEV (SAAS).....	9
3 SHRANJEVANJE PODATKOV V POSLOVNIH SISTEMIH	11
3.1 TIPI PODATKOV	11
3.1.1 <i>Transakcijski podatki</i>	11
3.1.2 <i>Dokumenti</i>	12
3.1.3 <i>Večpredstavnostni podatki (slike, video in glasba)</i>	12
3.2 PROBLEM SHRANJEVANJA PODATKOV NA LASTNI STROJNI OPREMI	12
3.3 EKONOMSKI IN PERFORMANČNI VIDIKI SHRANJEVANJA PODATKOV (TESTIRANJA).....	13
3.3.1 <i>Podatkovna baza v oblaku in na lastni strojni opremi</i>	14
3.3.1.1 <i>Performanse podatkovne baze na lastni strojni opremi</i>	15
3.3.1.2 <i>Performanse podatkovne baze v oblaku</i>	17
3.3.1.3 <i>Cenovna primerjava</i>	19
3.3.2 <i>Shranjevanje elektronskih dokumentov (fizično vs. oblak)</i>	22
3.3.2.1 <i>Shranjevanje v oblaku</i>	22
3.3.2.2 <i>Shranjevanje lokalno</i>	24
3.3.2.3 <i>Primerjava shranjevanja podatkov lokalno in v oblaku</i>	26
4 ZAKLJUČEK	29
KAZALO SLIK	31

KAZALO TABEL..... 32

LITERATURA 33

SEZNAM UPORABLJENIH KRATIC IN SIMBOLOV

- API *Application Programming Interface* (Aplikacijski programski vmesnik)
- CRM *Customer Relationship Management* (Sistemi za upravljanje odnosov s strankami)
- ERP *Enterprise resource planning* (Poslovni informacijski sistem)
- HRM *Human resource management* (Strategija upravljanja z zaposlenimi)
- IaaS *Infrastructure as a Service* (Infrastruktura kot storitev)
- PaaS *Platform as a Service* (Platforma kot storitev)
- RAID *Redundant Array of Inexpensive Disks* (Je standard povezovanja dveh ali več trdih diskov in upravljanja z njimi)
- PB Podatkovna baza
- PDF *Portable Document Format* (Odprt standard za izmenjavo elektronskih dokumentov)
- SaaS *Software as a Service* (Programska oprema kot storitev)
- SQL *Structured Query Language* (Poizvedovalni jezik za delo s podatkovnimi bazami)
- VM *Virtual Machine* (Navidezna naprava)

POVZETEK

Diplomska naloga govori o računalništvu v oblaku, danes zelo popularni temi. Njen cilj je bil ugotoviti, ali se podjetjem, iz ekonomskega in performančnega vidika, splača shranjevati podatke v oblaku.

V uvodu je tema na kratko predstavljena. V nadaljevanju naloga govori predvsem o uporabi storitev za shranjevanje podatkov, ki so zanimive za podjetja (pravne osebe). Podrobno opisuje računalništvo v oblaku in povzame kratko zgodovino njegovega razvoja. Opisuje tudi storitvene modele oblačnega računalništva. Nato predstavi različne podatke, ki jih lahko najdemo v okviru poslovnega sistema. Avtor se omejuje predvsem na transakcijske podatke in dokumente v elektronski obliki. Naredi testni program za testiranje hitrosti in odzivnosti podatkovne baze, ter napravi teste na lastni strojni opremi in v oblaku. Nadalje primerja različne ponudnike za shranjevanje podatkov in testira hitrosti njihovih storitev. Ponudnike analizira tudi glede na ekonomski vidik. Na koncu avtor še izpostavi najbolj primerne ponudnika iz ekonomskega in performančnega vidika.

V zaključnem delu diplomske naloge so predstavljene ključne ugotovitve o tem, katere podatke je koristno prenesti v oblak in zakaj. Avtor poda tudi svoje mnenje glede prenosa podatkov v oblak za določene poslovne sisteme.

Ključne besede

Računalništvo v oblaku, skupni stroški lastništva, podatki, dokument, ekonomski vidik, performančni vidik

ABSTRACT

The thesis treats today's very popular topic: cloud computing. The aim was to evaluate cost-effectiveness of cloud data storage as well as its performance in term of speed and throughput.

In the introduction the topic is briefly presented. The services which are important for small to medium size companies are discussed, especially services concerning data storage. Furthermore, the thesis describes in detail the cloud computing and summarizes brief history of its development. It also describes the cloud computing service models. Then a variety of information that can be stored in the cloud is presented, especially for small to medium size businesses. Author limits himself primarily to transaction data and documents in electronic form. A test program to measure the speed and responsiveness of the database was developed. Cost and performance for commercially available cloud solution was compared to local server solution. Second group of comparison was performed between some popular public domain storage solutions to local network storage. Cost, speed and data throughput were analyzed. At the end the author exposes the most appropriate provider from economic and performance perspective.

In the final part of the thesis the key findings on what one should have in mind when deciding between cloud and local solutions for data storage are presented. The author also gives his opinion on the transfer of data into the cloud for specific business systems.

Keywords

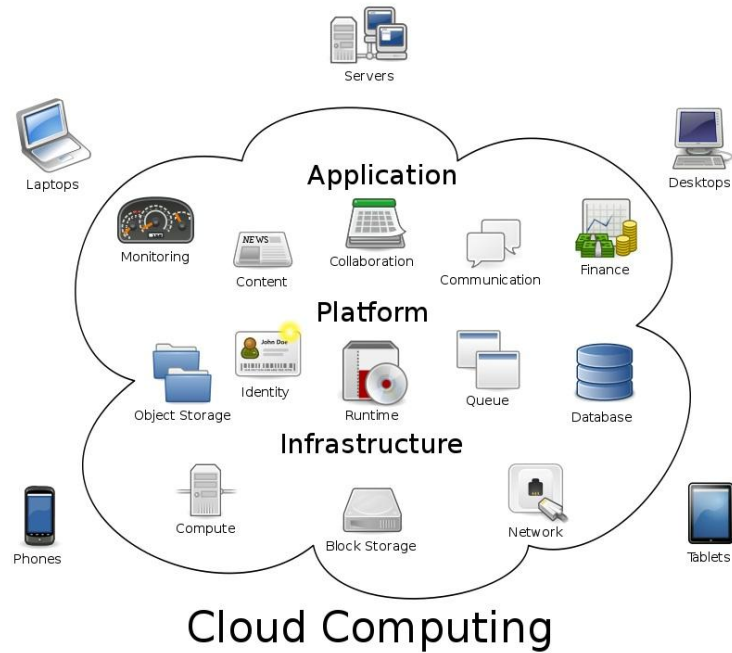
Cloud computing, total cost of ownership, data, document, economic aspect, performance aspect

1 UVOD

Računalništvo v oblaku je v današnjem času v naši družbi že močno razširjeno. Z njim se srečujemo tako rekoč vsakodnevno. Uporablja ga že velik del populacije. Uporabniki tako s pridom izkoriščamo različne storitve, ki jih tovrstno računalništvo ponuja. Primeroma naj navedem, da praktično vsaka fizična oseba, ki ima dostop do interneta, pozna in uporablja spletni iskalnik Google in njegove storitve. Ta je bil eden izmed prvih ponudnikov storitev v oblaku in je še danes med vodilnimi. Poleg njega so ponudniki tudi Yahoo, Hotmail,... Zelo veliko ljudi uporablja njihovo elektronsko pošto (gmail, hotmail, yahoo), nekateri pa uporabljajo tudi druge storitve (koledar, dokumente, slike, zemljevide,...).

Vse zgoraj naštetu velja predvsem za zasebne (fizične) uporabnike. Pri pravnih osebah je situacija nekoliko drugačna. Oblačno računalništvo bistveno spreminja koncept informacijske infrastrukture podjetji. Podjetje ne potrebuje več svojih strežnikov in aplikacij, zadostuje že osnovni računalnik, priključen v internetno omrežje. Vse aplikacije in podatki se nahajajo na nekam oddaljenem podatkovnem centru, pri ponudniku storitve. S tem si podjetja lahko bistveno znižajo stroške nakupa in vzdrževanja informacijske infrastrukture. Kljub temu računalništvo v oblaku trenutno uporabljajo v manjši meri kot fizične osebe. Navadno imajo svoje strežnike in vse njihove aplikacije (vključno z elektronsko pošto) praviloma tečejo na slednjih. Razlog za manjšo uporabo storitev računalništva v oblaku je, da tovrstnim storitvam še ne zaupajo popolnoma. Problem predstavlja dejstvo, da ne morejo pridobiti natančnih podatkov o tem, kje so shranjeni njihovi podatki in kdo vse lahko do njih dostopa in z njimi upravlja. Poleg varnosti je za večino podjetji problem tudi to, da bi morali vse svoje storitve preseliti v oblak (torej prenesti iz svojih strežnikov na strežnike ponudnikov storitev računalništva v oblaku), kar bi bilo seveda povezano s precejšnjimi stroški.

V svoji diplomski nalogi sem predstavil tiste storitve računalništva v oblaku, ki bi lahko bile zanimive za podjetja. Predstavil sem storitve za shranjevanje podatkov, večpredstavnostnih vsebin in dokumentov v elektronski obliki. Omejil sem se predvsem na transakcijske podatke in dokumente v elektronski obliki. Poskusil sem ugotoviti, ali se podjetjem splača prestaviti svoje poslovanje v oblak (ali bodo performance – predvsem hitrost, zadostile standardom podjetja) in koliko bi s tem prihranila.



Slika 1 Oblačno računalništvo [18].

2 RAČUNALNIŠTVO V OBLAKU (*angl. CLOUD COMPUTING*)

Izraz računalništvo v oblaku izhaja iz shematske upodobitve interneta, ki ga navadno uprizarjamo kot oblak (v preteklosti se je simbol oblaka uporabljal za upodobitev telefonskega omrežja). Oblak simbolizira zapleteno infrastrukturo, ki se nahaja nekje v ozadju. Nam se kaže preko veliko bolj prijaznega uporabniškega vmesnika. Oblak torej ponazarja nekaj, česar nam ni potrebno podrobno poznati.

Zametki oblačnega računalništva segajo v davno leto 1960. John McCarthy je že takrat napovedal, da bo računanje nekoč postalo organizirano tako kot del javne storitve (podobno kot ponudba električne energije). Do dejanske uresničitve njegove ideje je minilo kar nekaj desetletij. K temu je bistveno prispeval razvoj interneta. Včasih je bil internet enostranski medij, ker so uporabniki lahko zgolj pridobivali informacije iz spleta (lahko so samo brali različne članke in pridobivali informacije). Z razvojem pa je postal prava informacijska platforma. Danes lahko uporabniki preko spleta komunicirajo v vseh smereh, poleg suhoparnih besedil so se pojavile tudi gibljive slike (video), glasba in razne druge nove tehnologije. Poleg navedenega lahko danes preko spletnih brskalnikov zaganjamo in uporabljamo tudi različne (zapletene) aplikacije, katere tečejo na strežnikih nekje v »oblakih«.

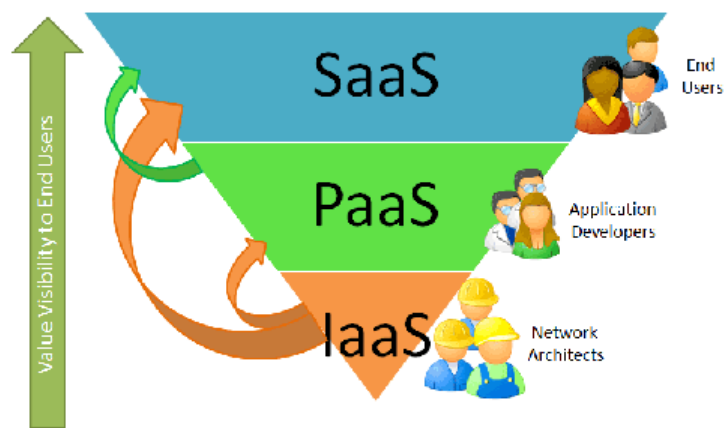
Računalništvo v oblaku (*angl. cloud computing*) je torej metafora, ki se uporablja za dinamično razširljiva, pogosto virtualizirana računalniška sredstva, ki so na voljo kot storitev preko interneta. Izraz oblak grafično ponazarja abstrakcijo tehnologije (računalniško omrežje ali internet). Poroča nam, da gre za nekaj, česar nam ni potrebno podrobno poznati. Računalništvo v oblaku pomeni, da so določeni viri strojne in programske opreme na različnih lokacijah, običajno zunaj lokacij uporabnika. Uporabniki pa glede na dejansko porabo najemajo, in tudi plačajo, potrebne informacijske vire. Poudariti je potrebno, da je bistvo oblačnega računalništva v tem, da uporabniki plačajo glede na dejansko porabo računalniških virov (procesorske moči in drugih virov). Uporabniku ni potrebno skrbeti za nakup, vzdrževanje strojne in programske opreme, ker za to poskrbijo ponudniki. Cilj ponudnikov pa je, da so njihovi viri čim bolj izkoriščeni (vsaj cca 95%), da si tako znižajo stroške [18, 14].

Poznamo 3 vrste oblakov, in sicer **javne**, **zasebne** in **hibridne** oblake. Javni oblaki so najbolj razširjena in poznana vrsta. So pod okriljem zunanjih organizacij, kot so Google, Yahoo, Amazon in drugi. Njihovi strežniki se nahajajo na različnih lokacijah po skoraj celem svetu in

so na voljo kateremu koli odjemalcu preko interneta. Problem javnih oblakov je predvsem varnost podatkov, saj naročniki ne vedo točno, kje so njihovi podatki shranjeni in kdo vse ima dostop do njih. Zasebni oblak je namenjen le enemu odjemalcu (podjetju). Najpogosteje so to oblaki znotraj velikih podjetji. Nadzor nad podatki ima to isto podjetje. Nivo varnosti je tako na najvišji ravni. Odjemalec si lasti vso opremo in ima posledično tudi popolni nadzor nad informacijskimi viri. Za vzpostavitev, upravljanje in vzdrževanje takšnega oblaka je potrebnih veliko strokovnjakov. Zaradi vsega naštetega je zasebni oblak veliko dražji od javnega. Tretja vrsta oblakov so hibridni oblaki. Ti združujejo javne in zasebne oblake. Notranji viri ostanejo pod nadzorom odjemalca (podjetja), za zunanje vire pa poskrbi ponudnik storitev. Tako so vsi občutljivi podatki shranjeni pri odjemalcu, podjetju pa je na voljo skoraj vsa skalabilnost javnega oblaka [11,23].

Računalništvo v oblaku lahko razdelimo na tri glavna področja oz. storitvene modele:

- Infrastruktura kot storitev (*angl. Infrastructure as a Service*) – IaaS,
- platforma kot storitev (*angl. Platform as a Service*) – PaaS,
- programska oprema kot storitev (*angl. Software as a Service*) – SaaS.



Slika 2 Storitveni modeli računalništva v oblaku [16].

Vsem trem področjem je skupno, da informacijske zmogljivosti zagotavlja ponudnik storitve. Uporabniki potrebujejo samo dostop do interneta in osnovni računalnik (ni potrebno da je zelo zmogljiv) [18, 8].

2.1 Infrastruktura kot storitev (IaaS)

IaaS je najbolj osnovni model oblačnega računalništva. Ponudnik storitve zagotovi računalnik (ali virtualno mašino), poskrbi za podatkovno shrambo in omrežje. Navadno imajo ponudniki več strežnikov, na katerih tečejo virtualne mašine. Vsak uporabnik dobi v uporabo eno virtualno mašino (računalnik), z v naprej zahtevanimi resursi. Uporabnik nato naloži željen operacijski sistem in ostalo programsko opremo. Potem vzpostavi povezavo z delovnim okoljem na daljavo. Uporabnik je sam odgovoren za vzdrževanje in delovanje operacijskega sistema in ostalih aplikacij. Ponudnik storitve zagotavlja in skrbi samo za strojno (virtualno) opremo in omrežje [18, 6].

Ta način zakupa strojne opreme je najcenejši izmed vseh treh področji, saj uporabnik plača samo dejansko porabo resursov (procesorja, delovnega pomnilnika in diskovja). V to ceno niso všteti stroški, povezani z nakupom programske opreme. Slabost tega sistema je, da mora uporabnik sam zagotoviti vzdrževanje in administracijo programskega dela sistema.

Največji in najbolj znani ponudnik tovrstnih infrastrukturnih oblačnih storitev je Amazonov Elastic Compute Cloud (EC2) [15].

2.2 Platforma kot storitev (PaaS)

Pri temu modelu oblačnega računalništva ponudnik zagotovi platformo, ki obsega operacijski sistem, določeno razvojno platformo, podatkovno bazo in spletni strežnik. Uporabniku ni potrebno skrbeti za strojno opremo, niti za programsko opremo (licence, instalacijo). Za vse to poskrbi ponudnik storitve. Tako se lahko uporabnik popolnoma posveti razvijanju svoje aplikacije [19].

Platformo kot storitev ponuja kar nekaj svetovno znanih podjetji. Najzanimivejši je Google, z njihovo storitvijo App Engine. Poleg njega sta znana tudi Microsoftov Azure Services Platform ter Amazonova EC2 in S3.

2.3 Programska oprema kot storitev (SaaS)

Programsko opremo kot storitev drugače imenujemo tudi storitev na zahtevo. To je model pri katerem ponudnik uporabniku nudi kompletno programsko in strojno opremo ter vse pripadajoče podatke. Kompletna programska (in strojna) oprema teče na ponudnikovih strežnikih v oblaku. Za dostop do različnih aplikacij, uporabnik potrebuje le osnovni računalnik, internetni brskalnik in dostop do interneta. Ponudniki uporabnikom omogočijo poganjanje aplikacij na njihovi infrastrukturi [21].

SaaS postaja vedno bolj popularna pri poslovnih aplikacijah npr. sistemi CRM (sistemi za upravljanje odnosov s strankami), ERP (poslovni informacijski sistem), HRM (strategija upravljanja z zaposlenimi). Poleg podjetji SaaS storitve uporablja tudi veliko zasebnih uporabnikov (npr. elektronsko pošto (Gmail, Yaho,...), dokumente (Google documents), zemljevide (Google Maps), slike (Picassa), itd...).

3 SHRANJEVANJE PODATKOV V POSLOVNIH SISTEMIH

Ključni del poslovnih sistemov predstavljajo podatki, ki so za vsako podjetje nenadomestljivi. To so transakcijski podatki (računi), dokumenti (podatki o naročilih, plačilih, zaposlenih, zalogah, partnerjih, strankah), večpredstavnostni podatki (video, slike, zvočne datoteke) in elektronska pošta, ki predstavlja vedno večjo količino podatkov. Količina podatkov, katere je potrebno shraniti, upravljati in arhivirati, v zadnjih letih nezadržno raste. Tudi vse napovedi različnih analitikov kažejo, da se bo količina podatkov v prihodnje samo še povečevala [7].

Podjetja imajo opravka z veliko količino podatkov, ki so lahko na različne načine hranjeni v podatkovnih skladiščih. Lahko so shranjeni v večjih podatkovnih bazah ali kot različni dokumenti na diskih. Veliko količino podatkov v podjetjih predstavljajo tudi različne varnostne kopije podatkov. Te so po navadi shranjene na magnetnih trakovih, saj je tak način shranjevanja podatkov cenejši in varnejši od shranjevanja na trdih diskih.

Količina teh podatkov, se v poslovnih sistemih iz leta v leto povečuje. Tako podjetja v parih letih poslovanja pridelajo enormne količine podatkov, katere morajo varno shranjevati, obdelovati in arhivirati. Pri tem se stalno povečujejo stroški za hrambo in procesiranje podatkov.

3.1 Tipi podatkov

Kot sem že omenil, se podjetja srečujejo z različnimi vrstami (tipi) podatkov. To so: transakcijski podatki, dokumenti, večpredstavnostni podatki (med katere uvrščamo slikovne podatke, video in glasbo),... V nadaljevanju sem bolj podrobno opisal posamezne tipe podatkov in njihovo rabo v podjetjih.

3.1.1 Transakcijski podatki

Transakcijski podatki opisujejo neke dogodke (notranji, zunanji) ali transakcije, ki se dogajajo v poslovnih procesih organizacij. Običajno so združeni v transakcijske zapise, ki vsebujejo podatke o času, numerične vrednosti in referenčne podatke. Informacijski zapisi o preteklih dogodkih, ki nastajajo kot interakcije med različnimi storitvami ali kupci, so transakcijske sledi [12, 22].

Izvor transakcijskih podatkov so poslovni dogodki, ki se beležijo v informacijskem sistemu neke organizacije. Med transakcijske podatke uvrščamo podatke o kupcih, zaposlenih, računih plačilih, zapisih dejavnosti, dobavah, skladiščnih evidencah, itd. Tipični sistemi, s katerimi obdelujemo transakcijske podatke so finančne aplikacije za celovite programske rešitve (ERP), sistemi za upravljanje odnosov s strankami (CRM),... Transakcijski podatki so navadno shranjeni v različnih podatkovnih bazah v podatkovnih skladiščih, zato jih ni težko obdelovati in analizirati. Z njihovo analizo lahko pridemo do informacij, znanja, vzorcev, nakupovalnih navad in pravil. To velja za transakcijske podatke, ki so služili za analizo. Razne druge (v mislih imam predvsem podatke o poslovanju, npr. računi) transakcijske podatke je potrebno shranjevati tudi zaradi pravil davčne in druge zakonodaje. Pri shranjevanju transakcijskih podatkov, ki vsebujejo osebne podatke zaposlenih ali drugih oseb, je treba biti še posebej previden. Upoštevati je potrebno določbe Zakona o varstvu osebnih podatkov (ZVOP-1; Ur.l. RS, št. [86/2004](#)).

3.1.2 Dokumenti

Pojem dokument je zelo širok. Zajema različne vrste dokumentov, ki se lahko pojavijo v različnih oblikah in predstavitev. Sam sem se v diplomski nalogi posvetil predvsem elektronskim dokumentom. Teh je v večjih organizacijah (podjetjih) ogromno, zato je pomembno, da so pametno in varno shranjeni (potrebno je izdelovati varnostne kopije dokumentov). Ti dokumenti se največkrat pojavljajo v obliki pdf (Portable Document Format), tekstovni dokumenti in drugi pisarniški dokumenti, ...

3.1.3 Večpredstavnostni podatki (slike, video in glasba)

Med večpredstavnostne podatke (*angl. multimedia*) uvrščamo slike, animacije, glasbo, video in kombinacijo vsega naštetega skupaj. V današnjih časih veliko količino podatkov predstavljajo prav tovrstni podatki.

3.2 Problem shranjevanja podatkov na lastni strojni opremi

Kot sem že omenil, nas v današnjih časih obdaja velika količina podatkov. Nič drugače ni pri podjetjih. Za shranjevanje večje količine podatkov je potreben primeren prostor. Prostor v katerih se nahajajo strežniki in podatkovna skladišča so navadno požarno in poplavno zaščiteni in primerno klimatizirani. Poleg tega je potrebno imeti zaposleno določeno število ustrezno usposobljenih ljudi, ki skrbijo za pravilno delovanje podatkovnih skladišč in strežnikov.

Zgoraj naštete stvari predstavljajo velik finančni zalogaj za podjetje. Zgraditi morajo primeren prostor, ga opremiti s strežniki in raznimi podatkovnimi napravami, zaposliti ustrezne ljudi, ki bodo skrbeli za nemoteno delovanje kompletnega računalniškega sistema v podjetju. K vsemu skupaj je potrebno prišteti še stroške programske opreme. Poleg vseh zgoraj naštetih stroškov ni zanemarljiva niti poraba električne energije, saj večina strežnikov in drugih komponent deluje 24 ur na dan, 7 dni v tednu. Z večanjem števila strežnikov, se poveča tudi število klimatskih naprav in njihova poraba električne energije.

Za manjša podjetja lahko takšen strošek predstavlja velik problem. Lahko se zgodi, da ne zmorejo financirati vsega zgoraj naštetega. S pojavom oblačnega računalništva, se odpirajo nove (cenejše) možnosti najema strojne in programske opreme. Mogoči so precej veliki prihranki. Predvsem manjšim podjetjem se zato izplača prestaviti informacijsko infrastrukturo v oblak.

V oblaku vsak uporabnik plača točno toliko kolikor porabi. Najame tudi toliko resursov (diskovnega prostora oz. procesorske moči) kolikor jih trenutno potrebuje. Ko se podjetje širi in potrebuje več prostora in močnejše resurse, zakupi večjo količino resursov. V primeru, da določenih resursov ne potrebujejo več, se temu primerno zmanjša znesek najema. Plačilo je torej odvisno zgolj od porabe sredstev.

Problem, ki se lahko pojavi pri oblačnem računalništvu, je hitrost prenosa podatkov med uporabnikom in oblakom. Vsi prenosi podatkov (ki jih ni malo) potekajo preko interneta. Zato ima ključno vlogo pri uporabnosti oblaka propustnost povezave z internetom.

V nadaljevanju sem prikazal hitrosti podatkovne baze v oblaku in jih primerjal s hitrostmi podatkovne baze na fizičnem računalniku. Primerjal sem tudi cene postavitve podatkovne baze na lastni strojni opremi in cene storitev v oblaku. Testiral sem hitrosti dostopa do različnih podatkovnih centrov v oblaku (Amazon CloudDrive, Google Drive, Dropbox) ter primerjal cene teh storitev.

3.3 Ekonomski in performančni vidiki shranjevanja podatkov (testiranja)

Zaradi pojava oblačnega računalništva je za shranjevanje podatkov mogoče uporabiti oblak. Na trgu je kar nekaj ponudnikov teh storitev. Ker je Amazon eden prvih in tudi največjih ponudnikov storitev računalništva v oblaku (ponuja tako raznovrstne aplikacije v oblaku, shranjevanje podatkov, arhiviranje, varnostne kopije in arhive,...), sem za cenovno primerjavo transakcijske podatkovne baze izbral prav njega.

Ker Amazon v času pisanja diplomske naloge ni imel preizkusne dobe svojih storitev (podatkovne baze), sem na internetu našel, drugega ponudnika Xeround (<http://xeround.com>). Ta ponuja 30 dnevni brezplačni preizkus njihove MySQL podatkovne baze v oblaku in uporablja Amazonove strežnike (Amazon EC2 storitev). Zato sem se registriral pri omenjenemu ponudniku, in tam tudi postavil relacijsko podatkovno bazo v oblaku. Enako bazo sem postavil tudi na svojem osebem računalniku, tako da sem lahko primerjal hitrosti in cene ene in druge storitve. Pri testiranjih sem se omejil na manjšo transakcijsko podatkovno bazo, ki simulira manjšo trgovino.

V drugem delu sem testiral tri ponudnike, ki ponujajo storitev za shranjevanje podatkov (dokumentov, videa, glasbe, slike,...). Izbral sem tri najbolj popularne ponudnike, in sicer Dropbox, Google Drive in Amazon Cloud Drive. Testiral sem hitrosti vsakega izmed njih ter primerjal njihovo ceno na GB, s ceno lastnega podatkovnega skladišča.

3.3.1 Podatkovna baza v oblaku in na lastni strojni opremi

V tem poglavju sem primerjal hitrosti in odzivnost relacijske podatkovne baze v oblaku in na fizičnem računalniku. Poleg primerjave hitrosti, sem naredil tudi ekonomsko primerjavo. Tako da, sem izračunal vse stroške obeh rešitev in jih primerjal med sabo.

Za testno okolje sem si izbral MySQL relacijsko podatkovno bazo (tako na fizičnem računalniku, kot v oblaku). Slednja teče na Linux operacijskem sistemu. Pri testiranju sem simuliral neko manjšo trgovino. Podatkovna baza se imenuje myTestDb, vsebuje pa 5 tabel: Uporabniki, Zaposleni, Artikli, Racun in PostavkeRacuna. Spodaj v tabelah 1 do 5 sem prikazal strukture posameznih tabel v bazi

ime	varchar(20)
priimek	varchar(20)
kraj	varchar(50)
postnaSt	char(4)
email	varchar(50)
idUp	varchar(20) PrimaryKey

Tabela 1 Struktura tabele Uporabniki.

ime	varchar(20)
priimek	varchar(20)
kraj	varchar(50)
postnaSt	char(4)
email	varchar(50)
idZap	varchar(20) PrimarKey

Tabela 2 Struktura tabele Zaposleni.

stRac	int(5)	
datumRac	date	
casRac	time	
placilnoSredstvo	char(4)	
znesekRac	int(6)	
popust	int(3)	
davek	int(6)	
idUp	varchar(20)	PrimayKey
idZap	varchar(20)	PrimayKey

Tabela 3 Struktura tabele Racun.

stArtikla	int(6)	PrimayKey
opisArtikla	varchar(50)	
cenaArtikla	int(6)	
davekArtikla	int(6)	

Tabela 4 Struktura tabele Artikli.

stRac	int(6)	PrimayKey
stArtikla	int(6)	PrimayKey

Tabela 5 Struktura tabele PostavkeRacuna.

Za vstavljanje v tabele in delo z njimi sem napisal testni program v programskem jeziku Java, v okolju Eclipse. Tako sem tabele Uporabniki in Zaposleni naključno napolnil tako, da je imela vsaka okoli 100 vrstic (imena, primke, kraje in poštne številke sem shranil v tekstovne datoteke in jih nato iz datotek naključno vstavljal v tabele, ostale vrednosti sem napolnil čisto naključno). Tabela Artikli sem napolnil z imeni naključnih artiklov iz interneta, tako da je vsebovala okoli 2500 vrstic. Tabeli Racun in PostavkeRacuna pa sem napolnil čisto naključno. Vsaka od njiju imela približno 1000 vrstic. Velikost podatkovne baze je znašala okoli 10 MB.

Ko sem imel napolnjene vse tabele (tako na fizičnem računalniku kot tudi v oblaku) in napisan testni program, sem se lotil testiranja hitrosti. To sem natančneje opisal v nadaljevanju, kjer sem v ločenih podpoglavjih predstavil rezultate testiranja vsake podatkovne baze posebej (lokalne in oblačne).

3.3.1.1 Performanse podatkovne baze na lastni strojni opremi

Za postavitev podatkovne baze sem uporabil svoj namizni računalnik na kateremu je nameščen 64-bitni operacijski sistem Windows 7. Specifikacije računalnika, na katerem sem

izvajal testiranja so bile: procesor Intel Core i5 750 (2,79GHz), delovni spomin 8GB, trdi disk 750GB in grafična kartica NVIDIA GeForce GTS250.

Uporabil sem MySQL podatkovno bazo (www.mysql.com), ki sem jo namestil na operacijski sistem Ubuntu 11.04. Ta je bil nameščen v virtualnem stroju Oracle VirtualBox (www.virtualbox.org). Nastavitve virtualnega stroja: dodelil sem mu dve jedri fizičnega procesorja, 3GB delovnega pomnilnika, 25GB trdega diska in 50MB grafičnega pomnilnika.

Za testiranje sem uporabil tudi prenosni računalnik, s katerim sem se povezoval na MySQL podatkovno bazo (ki je tekla na namiznem računalniku). Na njem je bil nameščen 32-bitni operacijski sistem Windows 7. Specifikacije prenosnega računalnika so bile naslednje: procesor Intel Core i5 U520 (1,07GHz), delovni spomin 4GB, trdi disk 320GB in Intelova grafična kartica, ki nima lastnega spomina.

Ko sem vse namestil sem se lotil testiranja. Strukturo podatkovne baze, ki sem jo uporabljal, je opisana že v točki 3.3.1., te diplomske naloge. V tem poglavju bom zato predstavil zgolj pridobljene rezultate. Testiral sem *vstavljanje*, *iskanje* in *brisanje* podatkov v tabelah. Pri testiranjih sem vedno vstavljaj, brisal ali iskal večje število vrstic (včasih tudi v zanki, tako da sem prišel do večjega števila ponovitev), potem pa sem izračunal povprečne čase. V tabelah sem predstavil rezultate, ki predstavljajo povprečne vrednosti več meritev. Vsi rezultati v tabelah se nanašajo na eno poizvedovanje (brisanje, vstavljanje ali iskanje) in so predstavljeni v mili sekundah (ms).

	Uporabniki	Zaposleni	Artikli
namizni računalnik	0,47 ms	0,45 ms	0,39 ms
prenosni računalnik ¹	1,91 ms	1,84 ms	1,65 ms

Tabela 6 Rezultati merjenja časa vstavljanja.

	Uporabniki	Zaposleni	Artikli
namizni računalnik	0,65 ms	0,59 ms	0,58 ms
prenosni računalnik ¹	2,25 ms	2,65 ms	2,12 ms

Tabela 7 Rezultati merjenja časa iskanja.

¹Prenosni računalnik se je na bazo povezoval preko brezžične povezave, katere hitrost je bila 36Mbps.

	Uporabniki	Zaposleni	Artikli
namizni računalnik	0,44 ms	0,48 ms	0,41 ms
prenosni računalnik ¹	1,95 ms	1,78 ms	1,74 ms

Tabela 8 Rezultati merjenja časa brisanja.

Meril sem tudi čas povezovanja na bazo (slika 3). To je čas, ki ga program porabi za vzpostavljane povezave in identifikacijo uporabnika.

```
Connection conn = DriverManager.getConnection( connection_url, "eclipse", "eclipse" );
Statement stmt = conn.createStatement();
```

Slika 3 Povezovanje na bazo.

Pri povezovanju namiznega računalnika na strežnik (ki je bil postavljen kot VM na istem računalniku) so časi za povezovanje na bazo znašali okoli 300ms do 350ms. Pri testiranju s prenosnim računalnikom, so bili časi višji. Gibali so se okoli 500ms do 600ms. Časi povezovanja na podatkovno bazo so bili neodvisni glede na vrsto in število poizvedovanja v bazi.

3.3.1.2 Performanse podatkovne baze v oblaku

Kot sem že omenil, sem si za testiranje podatkovne baze v oblaku izbral ponudnika Xeround (<http://xeround.com>), ki ponuja 30 dnevno brezplačno preizkušnjo, njegova storitev pa teče na Amazonovih strežnikih (Amazon EC2).

Strukturo podatkovne baze in testni program sem opisal na začetku tega poglavja. Na tej točki se bom tako bolj posvetil samim rezultatom. Omenim lahko le to, da sem se testiranja lotil na popolnoma enak način kot pri testiranju podatkovne baze na lokalnem (svojem) računalniku. Vse teste, ki sem jih opravil na domačem računalniku, sem opravil tudi v oblaku.

Pri testiranju hitrosti podatkovne baze v oblaku sem uporabljal prenosni računalnik (specifikacije so opisane že v 3. odstavku 3.3.1.1. poglavja). Testiranja sem opravljal na delovnem mestu, kjer imamo hitrejšo internetno povezavo, kot doma. Hitrost povezave je znašala 80Mbps² v smeri od ponudnika do firme in 40Mbps² v smeri prenosa iz službe do ponudnika internetnih storitev. Za primerjavo sem testiranja opravil tudi doma, na namiznem računalniku. Hitrost internetne povezave doma znaša 20/2Mbps³.

Tako kot pri testiranju na lokalnem (namiznem) računalniku, sem testiral *vstavljanje*, *brisanje* in *iskanje* podatkov v podatkovni bazi. Da so rezultati bolj verodostojni, sem vsako poizvedovanje v zanki opravil večkrat (ali pa sem napolnil celotno tabelo, cca 2500 vnosov).

² Testirano na strani www.speedtest.net. Dejanska linija je 100/50Mbps, vendar zaradi službenih procesov, kateri tečejo 24 ur na dan sem prišel do sledečih hitrosti.

³ Pri testiranju sem ugasnil vse ostale naprave in aplikacije, katere uporabljajo internet.

Potem sem izračunal povprečne vrednosti, ki so podane v spodnjih tabelah (tabela 9 do tabela 11). Vse vrednosti v tabelah se nanašajo na posamezno poizvedovanje (*angl.query*) in so izražene v milisekundah (ms).

	Uporabniki	Zaposleni	Artikli
prenosni računalnik ⁴	63,35 ms	67,04 ms	57,58 ms
namizni računalnik ⁵	72,53 ms	71,80 ms	67,55 ms

Tabela 9 Rezultati merjenja časa vstavljanja.

	Uporabniki	Zaposleni	Artikli
prenosni računalnik ⁴	89,12 ms	93,44 ms	71,05 ms
namizni računalnik ⁵	80,85 ms	82,66 ms	75,68 ms

Tabela 10 Rezultati merjenja časa iskanja.

	Uporabniki	Zaposleni	Artikli
prenosni računalnik ⁴	76,50 ms	72,45 ms	65,58 ms
namizni računalnik ⁵	75,98 ms	74,21 ms	70,74 ms

Tabela 11 Rezultati merjenja časa brisanja.

Čas povezovanja na bazo in identifikacija uporabnika je znašala od 980ms do 1200ms. Ugotovil sem, da je to bolj povezano z zasedenostjo ponudnika storitve kot pa s hitrostjo internetnega dostopa. Dogajalo se mi je namreč, da sem dosegal boljše čase povezovanja na bazo iz domačega računalnika (nižja hitrost linije), kot pa iz prenosnika v službi.

Tudi rezultati v tabelah (glej tabela 9 do tabela 11) so pokazali, da ni bistvene razlike med hitrostjo povezave 80/40Mbps in 20/2Mbps. V nekaterih primerih so razlike sicer vidne, po drugi strani pa se je včasih (izjemoma) tudi zgodilo, da so bili rezultati na hitrejši liniji, za malenkost slabši.

Kar se tiče občutka dela z bazo, se sicer pozna, da se baza nahaja v oblaku, vendar po mojem mnenju uporabnikov to ne bi smelo motiti pri delu, saj vseeno stvari potekajo tekoče. Opazil sem tudi, da ni skoraj nobene razlike v hitrosti internetnega dostopa med 20/2 Mbps in 80/40 Mbps linijo. Tudi ostale oblačne storitve (npr. igranje iger v oblaku), katere sem poizkušal doma (20/2 Mbps), so delale popolnoma brez problemov.

⁴ Testiranje opravljeno v službi, hitrost povezave 80/40Mbps.

⁵ Testiranje opravljeno doma, hitrost povezave 20/2Mbps.

3.3.1.3 Cenovna primerjava

Za cenovno primerjavo infrastrukture oz. storitev za postavitev podatkovne baze, sem izbral HP strežnik Proliant ML150G6, za ponudnika oblačnih storitev pa Amazon RDS (*angl. Relation Database Service*). V nadaljevanju sem predstavil obe rešitvi.

Za **postavitev podatkovne baze na fizičnem računalniku** sem izbral HP Proliant ML150G6 strežnik. Na njem sem namestil odprtokodni operacijski sistem Linux (Ubuntu). Za podatkovno bazo sem uporabil MySQL. Za tako konfiguracijo sem se odločil predvsem zaradi cene, saj sta tako Linux operacijski sistem kot tudi programska oprema MySQL, brezplačna. Poleg tega mislim, da ta HP strežnik zadostuje za potrebe neke manjše trgovine. Z dokupom majhnega števila komponent bi bil verjetno zmožen izvajati tudi mnogo zahtevnejša opravila. Poleg tega ima strežnik 3 letno garancijsko dobo.

Specifikacije in cena za Strežnik HP Proliant ML150G6 E5506 470065-43: [4]

procesor Intel® Xeon® E5506

(2.13GHz/4-core/4MB/80W)

predpomnilnik 4MB L3

pomnilnik 2 GB (1 x 2GB) 2Rx8 PC3-10600R-9

krmilnik HP SA P410/ZM (RAID 0,1,1+0)

trdi disk 300 GB SAS LFF Hot Plug

SATA DVD/RW

mreža HP NC107i PCI Express Gigabit

napajalnik 750W CS HE Hot Plug

ohišje 5U Tower, miška

garancija 36 mesecev (3-3-3)

Cena: **1.364,88€** (akcijska: 1.172,46€) [4]

Predpostavil sem, da amortizacijska doba za strežnik znaša 36 mesecev (3 leta). Mesečni strošek strežnika tako znese približno **38€**. Potrebno je poudariti, da je vsa programska oprema brezplačna, plačati je potrebno le strojno opremo. V primeru, da bi izbral katero drugo rešitev, bi bilo potrebno kupiti tudi programsko opremo. Njena cena bi predstavljala precejšen strošek. V ceno je potrebno všteti še vzdrževanje in nameščanje programske opreme. Potrebujemo tudi primeren prostor za strežnik in ostalo informacijsko opremo. Pri primerjavi sem izpustil strošek električne energije, saj je ta verjetno dokaj nizek.

Za **shranjevanje podatkovne baze v oblaku** sem izbral Amazonovo rešitev Amazon RDS (relacijska podatkovna baza). Amazon ima v ponudbi tri različne podatkovne baze: MySQL, Oracle in Microsoft SQL. Izbral sem si MySQL podatkovno bazo, ker je najcenejša in ker sem jo uporabljal že pri prejšnjih primerjavah [1].

Ker gre za relacijsko podatkovno bazo, lahko svoje obstoječe aplikacije, ki se nahajajo na našem fizičnem strežniku, lažje prenesemo v oblak. Ravno tako lahko še vedno dostopamo do te podatkovne baze (ki je sedaj v oblaku) z enakim orodjem kot prej. Instance podatkovne

baze lahko ustvarimo preko administracijske konzole ali preko vmesnika API (*angl. Application programming interface*).

Poleg podatkovne baze, je potrebno izbrati še velikost instance (podatkovne baze). Izbiramo lahko med sedmimi instancami (od mikro instanc, do štirikratnih zelo velikih instanc). Za primerjavo, v diplomski nalogi sem izbral »majhno« instanco (*angl. small DB instance*). To dejansko pomeni, da zakupimo eno virtualno jedro procesorja s hitrostjo 1 ECU⁶ (*angl. elastic compute units*), dodeljenega nam je 1.7 GB delovnega spomina, vse skupaj pa teče na 64-bitni platformi. V primeru drugačnih potreb lahko kadarkoli zakupimo drugačno (zmogljivejšo) konfiguracijo računalniških virov.

Pri najemu podatkovne baze, lahko izberemo še regijo, kje naj bodo shranjeni naši podatki. Možnosti so ZDA, Evropa, Azija in Južna Amerika. Glede na regijo se spreminja tudi cena najema (najceneje je ZDA, najdražja pa Južna Amerika). Sam sem izbral Evropo in sicer podatkovni center na Irskem (edina možnost za Evropo). Cena mesečnega najema je odvisna še od velikosti podatkovne baze, števila vhodno / izhodnih operacij in prenosa podatkov preko interneta. V tabelah 12 do 14 so navedene cene.

Instanca Podatkovne Baze	Cena na uro
Mikro instanca (<i>angl. micro DB Instance</i>)	0.035 \$
Majhna instanca (<i>angl. small DB Instance</i>)	0.115 \$
Velika instanca (<i>angl. large DB Instance</i>)	0.455 \$
Zelo velika (<i>angl. extra Large DB Instance</i>)	0.920 \$

Tabela 12 Cene za posamezne instance (Evropska regija) [1].

Shramba podatkov	0.11 \$ na GB na mesec
Vhodno / izhodni dostopi	0.11 \$ na 1 mio.zahtevkov

Tabela 13 Cene za hrambo podatkov in vhodno / izhodne dostope [1].

Poleg vsega naštetega Amazon zagotavlja tudi avtomatično varnostno kopiranje podatkovne baze, tako da lahko v vsakem trenutku povrnemo izgubljene podatke ali celo obnovimo stanje podatkovne baze na določen dan. Iz administratorske konzole je jasno in natančno razvidno koliko in katere vire smo že porabili, kakšno konfiguracijo imamo izbrano, aktivnost vhodno / izhodnih dostopov in število povezav posameznih instanc podatkovne baze.

⁶ 1 ECU je ekvivalenten hitrosti procesorja Opteron (2007) ali Xeon (2007) s hitrostjo med 1.0 – 1.2 GHz

Prenos podatkov do Amazona	Cena
Brezplačno	/
Prenos podatkov od Amazona	
Prvi 1 GB / mesec	0.0 \$ na GB
Do 10 TB / mesec	0.12 \$ na GB
Do 40 TB / mesec	0.09 \$ na GB
Do 100 TB / mesec	0.07 \$ na GB
Do 350 TB / mesec	0.05 \$ na GB

Tabela 14 Cena za prenose podatkov (Evropska regija) [1].

Glede na mojo testno podatkovno bazo, opisano v poglavju 3, sem izračunal mesečne stroške najema podatkovne baze pri Amazonu. Predpostavil sem, da je podatkovna baza aktivna 8 ur na dan, 5 dni v tednu. Iz tega izhaja, da se podatkovno bazo uporablja približno **180 ur** na mesec (ob nadaljnji predpostavki, da mesec sestavljajo 4 tedni in pol). Velikost same baze (okoli 10 MB) na mesečne stroške skoraj nima vpliva, saj je obračunski interval 1 GB. Prav tako je zanemarljiv strošek shrambe podatkov ter vhodno / izhodni dostopi (našteto sem vseeno upošteval pri končnem izračunu in na ta način potrdil tezo, da to skoraj ne vpliva na ceno). Za prenos podatkov iz Amazonovih strežnikov sem vzel vrednost 10 GB na mesec (po vsej verjetnosti sem malo pretiraval).

Izračun stroškov:

$$180 \text{ ur} * 0,115\$ (\text{uporaba baze}) + 0,11\$ (\text{shramba podatkov}) + 0,11\$ (V / I \text{ dostopi}) + 9 * 0,12\$ (\text{prenosi podatkov}) = 22 \$ = \mathbf{18,05 \text{ €}^7}$$

V primeru gostovanja MySQL podatkovne baze na Amazonovih strežnikih bi za v 3. poglavju opisano podatkovno bazo, morali odšteti slabih 20€ na mesec.

Glede na oba izračuna podatkovne baze (v oblaku in na fizičnem računalniku) sem ugotovil, da je bistveno ceneje postaviti podatkovno bazo v oblaku. Razlika med obema primeroma znaša 20€ na mesec, kar ni zanemarljiv strošek. Ta razlika bi se krepko povečala, če bi pri postavitvi podatkovne baze (na fizičnem računalniku) izbral katero drugo (plačljivo) programsko opremo in upošteval strošek vzdrževanja. Poleg tega, je baza v oblaku dosti bolj prilagodljiva, saj v primeru, da potrebujemo zmogljivejšo bazo, samo najamemo zmogljivejše računalniške vire. Poleg tega nam ni potrebno skrbeti za vzdrževanje in za varnostno kopiranje, kar je pri lastni strojni opremi potrebno.

⁷ Menjalni tečaj 1 EUR = 1,22 USD (2.8.2012)

3.3.2 Shranjevanje elektronskih dokumentov (fizično vs. oblak)

V tem poglavju sem primerjal shranjevanje podatkov v oblaku in lokalno. Zanimalo me je, če je oblak smiselno uporabiti za shranjevanje podatkov v elektronski obliki. Predvideval sem, da je iz performančnega vidika takšna smiselnost najverjetneje podana. Kljub temu sem testiral hitrosti posameznih ponudnikov. Proučil sem tudi ekonomsko plat. Izračunal sem stroške obeh variant (oblaka in lokalno) in nato napravil zaključek, katera od obeh variant je bolj primerna.

3.3.2.1 Shranjevanje v oblaku

Za shranjevanje različnih dokumentov v oblaku imamo kar veliko ponudnikov. Odločil sem se, da bom opisal in testiral tri izmed njih. Eden prvih ponudnikov na trgu je bil Dropbox [3]. Ta fizičnim uporabnikom brezplačno ponuja od 2GB do 18GB prostora na njihovih diskih. Kmalu za njim sta svoji storitvi ponudila tudi dva velika igralca na področju oblachnega računalništva, Amazon in Google. Prvi ponuja svojo storitev kot Amazon Cloud Drive [2] in uporabnikom ponuja 5GB brezplačnega prostora. Kot zadnji izmed izbranih treh ponudnikov, je svojo storitev predstavil Google, imenuje se Google Drive [5]. Tudi oni ponujajo 5GB brezplačnega prostora, s tem da se Googlovi dokumenti ne štejejo v kvoto 5GB (štejejo se samo dokumenti, katere uporabnik naloži iz svojega računalnika).

Testiranja hitrosti prenosa podatkov sem se lotil tako, da sem kopiral video datoteko iz svojega računalnika (prenosnika) na vsakega izmed izbranih treh ponudnikov (in seveda tudi v obratni smeri, od ponudnika na svoj prenosnik). Hitrost povezave je bila 80/40 Mbps. Velikost videa je bila približno 500MB. Med prenašanjem filma v oblak, sem imel vse ostale stvari na računalniku izklopljene, da nebi po nepotrebnem »zasedale« pasovne širine. Pri Dropboxu sem za kopiranje filma uporabil kar njihov program (Dropbox ga prikaže kot eno izmed map na računalniku). Pri Google Drive in Amazon Cloud Drive pa sem kopiral preko web vmesnika. Preko programov, ponujenih iz njune strani, se namreč ni dalo učinkovito meriti časa prenosa. Pri tem še pripominjam, da ima po mojem mnenju Dropbox najboljši program za kopiranje datotek. Slednji izgleda kot mapa na disku in se ves čas sinhronizira z oblakom. V tabeli 15 so prikazani rezultati vseh treh ponudnikov.

	download [MB/s]	upload [KB/s]
Dropbox	1,89	901,8
Amazon Cloud Drive	1,53	720,4
Google Drive	4,20	975,5

Tabela 15 Rezultati testiranja hitrosti prenosa.⁸

⁸ Prenašal sem 500MB velik film, hitrost povezave je bila pa 80/40Mbps.

Kot je razvidno iz tabele 15 je Google Drive najhitrejši v smeri od ponudnika k uporabniku in v obratni smeri. Opazil sem sicer, da pri njem hitrosti nekoliko variirajo, kljub temu pa je bil vedno najhitrejši. Dropbox ima zelo konstante hitrosti prenosov. Prenos na njihove strežnike je vedno okoli 900 KB/s, prenos v obratni smeri se pa giblje med 1 MB/s do 2 MB/s. Amazon se je tokrat najslabše izkazal. Imel sem probleme pri prenašanju nekaterih dokumentov v pdf formatu. Določen dokument je bil sprejet šele po večkratnih poskusih prenosa. Pri prvem poskusu mi je sistem sporočil, da je prišlo do napake in prenos ni bil uspešen. To se je dogajalo kar pogosto, še posebej če sem kopiral večje število pdf dokumentov.

Kar se tiče same hitrosti in uporabe shranjevanja podatkov v oblaku sta v ospredju Google Drive in Dropbox, Amazonov Cloud Drive me ni prepričal. Prednost Googla je hitrost prenosa podatkov v obe smeri. Na drugi strani ima Dropbox boljše dodelam program, saj se popolnoma integrira v operacijski sistem, tako da zglada kot del trdega diska na računalniku. Tudi sinhronizacija deluje odlično, posebno med več različnimi računalniki oz. drugimi (mobilnimi) napravami.

prostor	cena / mesec	cena / leto
2 GB (do 18 GB)	brezplačno	brezplačno
50 GB (do +32 GB)	9,99 \$	99 \$
100 GB (do +32 GB)	19,99 \$	199 \$

Tabela 16 Dropbox cenik [3].

Poleg performančnih značilnosti, sem pregledal tudi ekonomski vidik shranjevanja podatkov v oblaku. Vsi trije ponudniki ponujajo nekaj GB prostora brezplačno. Največ brezplačnega prostora dobimo pri Googlu in Amazonu, 5 GB. Pri Dropboxu ima vsak uporabnik na začetku na voljo samo 2 GB, vendar si jih lahko prisluži še dodatnih 16 GB, tako da privede nove člane (novi član se mora registrirati in inštalirati njihov program na svoj računalnik).

prostor	cena / leto
5 GB	brezplačno
20 GB	20 \$
50 GB	50 \$
100 GB	100 \$
200 GB	200 \$
500 GB	500 \$
1000 GB	1000 \$

Tabela 17 Amazon Cloud Drive cenik [2].

Vsi trije ponudniki ponujajo več prostora v oblaku za določeno vsoto denarja. Tukaj je Google najcenejši in Dropbox najdražji. Pri slednjem je potrebno odšteti 199 \$ na leto, za 100

GB prostora (s privedbo novih članov se lahko dobi še do 32 GB dodatnega prostora) kar je razvidno iz tabele 16.

Pri Amazonu bomo za 100 GB spomina odšteli 100 \$ na leto. Cena za vsak nadaljnji GB zraste za 1 \$, kar je navedeno tudi v tabeli 17. Pri zadnjem, najcenejšem ponudniku (Googlu) bomo za 100 GB prostora plačali le 4,99 \$ na mesec oziroma 59,88 \$ na leto. Ostale cene tudi tukaj naraščajo linearno in so podane v tabeli 18.

prostor	cena / mesec	cena / leto ⁹
5 GB	brezplačno	brezplačno
25 GB	2,49 \$	29,88 \$
100 GB	4,99 \$	59,88 \$
200 GB	9,99 \$	119,88 \$
400 GB	19,99 \$	239,88 \$
1 TB	49,99 \$	599,88 \$
2 TB	99,99 \$	1199,88 \$
4 TB	199,99 \$	2399,88 \$
8 TB	399,99 \$	4799,88 \$
16 TB	799,99 \$	9599,88 \$

Tabela 18 Google Drive cenik [5].

Ker je Googlov Drive najcenejši in tudi performanično najboljši, sem ga izbral za primerjavo z lokalnim shranjevanjem podatkov.

3.3.2.2 Shranjevanje lokalno

Za shranjevanje podatkov na lastni strojni opremi potrebujemo strežnik NAS (*angl. network attached storage*) oziroma strežnik na katerega priključimo ustrezne diske. Potrebujemo tudi posebni prostor. V primeru, da imamo veliko številko diskov in strežnikov, mora biti prostor tudi ustrezno klimatiziran. V diplomski nalogi sem se omejil na neko mini podatkovno skladišče, tako da posebni prostor in klimatizacija nista nujno potrebna.

Za podatkovno skladišče sem si izbral strežnik NAS. To je mini računalnik na katerega je priključenih več trdih diskov, ki so uporabnikom dostopni preko lokalnega omrežja. NAS je čisto samostojna enota, na katero priklapljam diske. Ti so nameščeni v različna polja RAID (*angl. redundant array of independent disks*). To je standard povezovanja dveh ali več trdih diskov in upravljanja z njimi. Nastal je z namenom, da bi lahko več manjših in počasnejših diskov, povezali v večjo, hitrejšo in bolj zanesljivo logično enoto. Poznamo več načinov RAID povezovanja. Lahko nam omogoča izboljšano zmogljivost, dodaten prostor ali oboje

⁹ Preračunano glede na ceno / mesec pomnoženo z 12 mesecev. Google ima v ceniku samo ceno na mesec.

hkrati. Za večjo zanesljivost, potrebujemo več diskov za manjši prostor (podatki se podvajajo), kot pri manjši zanesljivosti. [20]

Poleg NAS enote potrebujemo še trde diske. V svojem primeru (simuliral sem manjše podatkovno skladišče) sem izbral 5 trdih diskov (WD 2003FYYS 2TB RaidEdition). Vsak ima 2 TB prostora. Preko NAS enote Thecus N4100EVO so povezani v polje RAID 5, kar pomeni da so štirje diski uporabni, peti disk je pa varovalni disk. V primeru okvare enega izmed diskov so podatki še vedno varno shranjeni na ostalih. Če se pokvarita dva diska hkrati, pa del podatkov izgubimo.

Specifikacije in cena enota NAS Thecus N4100EVO: [9]

možnost vgradnje: 5x 3.5"/2.5" SATA

RAID nivoji: 0/1/5 in JBOD

procesor: Cavium ECONA CNS3420 SoC 600MHz (Dual-Core)

pomnilnik: 256MB DDR

mrežni žični vmesnik: 2x 1Gbps

USB vhodi: 2x USB 2.0

dimenzije: 226 x 165 x 192 mm

teža: 5kg

poraba el .energije: v spanju 60W / v delovanju 80W (odvisno od kombinacije)

Cena: 398,66€ (v akciji pa 378,73€)

Specifikacije in cena trdega diska WD 2003FYYS 2TB RaidEdition: [10]

družina: Caviar RE4 (Raid Edition)

neprekinjeno delovanje 24/7: Da

model: WD2003FYYS

kapaciteta: 2TB

tip vmesnika: SATA300

hitrost vrtenja plošč: 7.200 obr/min

velikost predpomnilnika: 64MB

NCQ: Da

tekočinski ležaji: Da

Cena: 243,00€ (v akciji pa 230,86€)

Torej strošek strojne opreme za podatkovno skladišče 8 TB podatkov znesse **1.613,66€** (5*243€ + 398,66€). To pomeni **0,202€** (1613,66€ / 8000 GB) na GB. Določil sem, da amortizacijska doba znaša 5 let (60 mesecev). Mesečni strošek za strojno opremo tako nanese **26,89€**. V našem primeru to velja za 8 TB veliko podatkovno skladišče. Ker bo naše podatkovno skladišče dostopno 24 ur na dan 7 dni v tednu, ni zanemarljiv podatek o porabi električne energije. Glede na specifikacije naj bi enota NAS porabila 60W energije v mirovanju in 80W pri delovanju. Za izračun sem predpostavil povprečno porabo 70W. To na mesec znesse dobrih **50kWh** (0,07W*24h*30 dni). Glede na trenutno ceno električne energije, ki znaša 0,06551€ na kWh pri Petrolu, [13] bi mesečni strošek znašal dobre **3,3 €** (50kWh*0,06551€/kWh). V tej ceni je upoštevana samo cena električne energije, ni pa

omrežnine in drugih stroškov. Cena podatkovnega skladišča tako naraste na nekaj čez **30 €** mesečno.

3.3.2.3 Primerjava shranjevanja podatkov lokalno in v oblaku

V tem delu svoje diplomske naloge sem se osredotočil na prednosti in slabosti shranjevanja podatkov v oblaku in na lastni strojni opremi. Prednost slednje je predvsem hitrost prenosa podatkov. V lokalnem omrežju je namreč ta nekajkrat večja kot v oblaku. Dejstvo pa je, da to za potrebe neke majhne trgovine ne igra pomembne vloge. Tudi hitrosti, ki sem jih dosegel pri testiranju prenosov med oblakom in lokalnim računalnikom (prenos na oblak malo pod 1 MB/s in prenos iz oblaka okoli 4 MB/s) so namreč zelo solidne. Pri prenašanju raznih elektronskih dokumentov (pod te dokumente uvrščam predvsem razne dokumente PDF, Word, itd...) ne bo nobenih problemov. Malo dalj časa bi trajal prenos večjih video datotek, katerih pa v neki manjši trgovini ni veliko (v kolikor bi se kakšen le našel, bi ga lahko pustili, da bi se prenesel čez noč).

Slabost lastne strojne opreme je predvsem vzdrževanje. V primeru, da gre kaj narobe, namreč potrebujemo strokovnjaka, ki bo popravil (zamenjal) trdi disk in rešil podatke. Pri oblaku teh problemov nimamo. Za vse to poskrbi ponudnik storitve. Tudi možnost izgube podatkov, je v oblaku manjša kot pri lokalnem shranjevanju. V oblaku so podatki shranjeni v različnih podatkovnih centrih razporejenih po celem svetu, medtem ko imamo pri lastni strojni opremi vse podatke na enem mestu. V primeru neke večje neravne nesreče (poplave, požar,...) tako izgubimo vse podatke. Večja podjetja imajo za to varnostne kopije svojih podatkov na drugi lokaciji kot je podjetje, to pa pomeni, da potrebujejo dva velika podatkovna centra. Oba je potrebno zgraditi protipožarno in protipoplavno, ju vzdrževati, ... Skratka stroški se s tem zelo povečajo.

Kar se tiče varnosti podatkov, so nekateri zelo skeptični do Googla. Večino podatkov, ki so shranjeni na njihovih strežnikih pregledujejo s svojimi algoritmi, lahko pa jih celo uporabijo za oglaševalske namene. Podatki se od ponudnika k nam prenašajo preko varne https povezave, tako da ni tako lahko pridobiti tuje podatke. Po drugi strani, pa lahko na lokacijo kjer je podatkovno skladišče na lokalni strojni opremi, tudi udrejo nepridipravi in dobijo vse naše podatke.

Največje vprašanje je ali se nam splača shranjevati podatke na lastni strojni opremi, ali je cenejši oblak (torej ekonomski vidik uporabe storitev v oblaku). Če primerjam zgolj svoje izračune, torej ceno za 8 TB podatkov na lastni strojni opremi in v oblaku, se izkaže, da je shranjevanje na lastni opremi cenejša rešitev. Pri tem je potrebno poudariti, da pri shranjevanju v oblaku plačamo toliko kolikor v tistem trenutku potrebujemo. Ni nam namreč potrebno najeti celih 8 TB prostora že na samem začetku. Tako nam na primer Google ponuja najem 25 GB za ceno približno 2 € na mesec. S časoma, ko potrebujemo več prostora lahko

tega, enostavno najamemo in tudi temu primerno več plačamo. Poleg tega nam ni potrebno skrbeti za vzdrževanje, varnostno kopiranje itd. Pri lastnem podatkovnem centru je večji strošek električne energije. Poleg tega ne moremo vsak mesec dokupovati novega prostora, oziroma se nam ne splača postaviti podatkovni center s samo 100 GB prostora.

Glede na vse navedeno se po mojem mnenju za manjše količine prostora definitivno bolj izplača najeti nekega ponudnika v oblaku (sam se najbolj nagibam k Google-u), saj za zelo nizko ceno ponujajo kar nekaj prostora in kompletno vzdrževanje skupaj z varnostnimi kopijami.

4 ZAKLJUČEK

V diplomski nalogi sem proučil uporabnost računalništva v oblaku za poslovanje pravnih oseb. Najprej sem na splošno opisal kaj je računalništvo v oblaku in na kratko predstavil področja njegove uporabe (storitvene modele). Ta področja so infrastruktura kot storitev, platforma kot storitev in programska oprema kot storitev. Storitve, s katerimi sem se ukvarjal v nadaljevanju, spadajo v področji platforme (baze podatkov) in programske opreme (shranjevanje elektronskih dokumentov) kot storitve. Obdelal sem storitve shranjevanja podatkov (relacijske podatkovne baze) v poslovnih sistemih, storitve večpredstavnostnih vsebin in dokumentov v elektronski obliki. Predstavil sem shranjevanje podatkov v oblaku ter shranjevanje na lastni strojni opremi in opisal prednosti in slabosti ene in druge oblike.

Ugotovil sem, da je transakcijske podatke (podatkovne baze) vsekakor koristno prenesti v oblak, tako iz ekonomskega kot tudi iz performančnega vidika. Odzivnost je sicer nekoliko slabša kot pri shranjevanju na lastni strojni opremi (rezultati v tabelah 6 do 11), vendar to, po mojem mnenju, uporabnikom ne povzroča nikakršnih težav. Glede shranjevanja elektronskih dokumentov menim, da je prenos v oblak manj smiseln. Iz ekonomskega vidika je prenos koristen pri manjših količinah podatkov (do okoli 500 GB). Pri večjih količinah je finančno bolj ugodno podatke shraniti na lastni strojni opremi (cena za shranitev 8 TB podatkov v oblaku je približno 330 EUR na mesec, medtem ko cena za shrambo enake količine podatkov v podatkovnem skladišču znaša približno 30 EUR mesečno). Iz performančnega vidika se, po mojem mnenju, shranjevanje elektronskih dokumentov v oblaku ne izplača, ker je prenos (predvsem večjih večpredstavnostnih vsebin) veliko počasnejši (podatki v tabeli 15). Za manjše količine podatkov (posamične elektronske dokumente) je shranjevanje v oblaku tudi iz tega vidika smiselno. Razlika v hitrosti, zaradi majhnega časa trajanja prenosa, v tem primeru ni moteča.

Ponudniki računalništva v oblaku zagotavljajo varnost tega načina shranjevanja podatkov. Klub temu problem varnosti ostaja odprt. Uporabniki namreč nikoli točno ne vedo, kje se njihovi podatki nahajajo in kdo vse lahko do njih dostopa. Ta problem v določeni meri rešujejo hibridni oblaki. V primeru slednjih ima pravna oseba svoje podatkovno skladišče, ostale računalniške vire (procesorska moč, delovni spomin, ipd.) pa najame od ponudnikov oblachnega računalništva. Problem, zaradi katerega se (predvsem večja) podjetja še ne odločajo za shranjevanje podatkov v oblaku, je tudi ne-poenotena zakonodaja (razlike med zakonodajo držav in evropsko zakonodajo).

Glede na vse navedeno sem prišel do zaključka, da se manjšim podjetjem vsekakor splača uporabljati storitve računalništva v oblaku. Na ta način si močno znižajo začetne stroške poslovanja (za nakup strojne in programske opreme ter za samo administracijo in vzdrževanje neke platforme). Za večja podjetja je računalništvo v oblaku manj koristno, zato se, po mojem mnenju, zanj (še) ne bodo odločala. Problem so zelo velike količine podatkov, določene zakonske omejitve, varnostni vidiki in podobno. Tudi sama selitev podatkov iz lastnih baz v oblak bi bila zamudna in draga.

KAZALO SLIK

SLIKA 1 OBLAČNO RAČUNALNIŠTVO [18].....	6
SLIKA 2 STORITVENI MODELI RAČUNALNIŠTVA V OBLAKU [16].	8
SLIKA 3 POVEZOVANJE NA BAZO.	17

KAZALO TABEL

TABELA 1 STRUKTURA TABELE UPORABNIKI.....	14
TABELA 2 STRUKTURA TABELE ZAPOSLENI.....	14
TABELA 3 STRUKTURA TABELE RACUN.....	15
TABELA 4 STRUKTURA TABELE ARTIKLI.....	15
TABELA 5 STRUKTURA TABELE POSTAVKE RACUNA.....	15
TABELA 6 REZULTATI MERJENJA ČASA VSTAVLJANJA.....	16
TABELA 7 REZULTATI MERJENJA ČASA ISKANJA.....	16
TABELA 8 REZULTATI MERJENJA ČASA BRISANJA.....	17
TABELA 9 REZULTATI MERJENJA ČASA VSTAVLJANJA.....	18
TABELA 10 REZULTATI MERJENJA ČASA ISKANJA.....	18
TABELA 11 REZULTATI MERJENJA ČASA BRISANJA.....	18
TABELA 12 CENE ZA POSAMEZNE INSTANCE (EVROPSKA REGIJA) [1].....	20
TABELA 13 CENE ZA HRAMBO PODATKOV IN VHODNO / IZHODNE DOSTOPE [1].....	20
TABELA 14 CENA ZA PRENOSE PODATKOV (EVROPSKA REGIJA) [1].....	21
TABELA 15 REZULTATI TESTIRANJA HITROSTI PRENOSA.....	22
TABELA 16 DROPBOX CENIK [3].....	23
TABELA 17 AMAZON CLOUD DRIVE CENIK [2].....	23

LITERATURA

- [1] (2012) Amazon web service. »Amazon Relational Database Service.« Dostopno na: <http://aws.amazon.com/rds/>
- [2] (2012) Amazon. »Learn More About Amazon Cloud Drive.« Dostopno na: http://www.amazon.com/gp/feature.html?ie=UTF8&nav_sdd=aps&docId=1000796931&pf_rd_m=ATVPDKIKX0DER&pf_rd_s=center-A5&pf_rd_r=0FQJ4EBD2JJ3A8AN0NGA&pf_rd_t=101&pf_rd_p=1364264882&pf_rd_i=507846&ref_=cd_lm_rd_fp
- [3] (2012) Dropbox. »Dropbox – Simplify your life.« Dostopno na: <http://www.dropbox.com>
- [4] (2012) EnaA.com. »HP strežniki.« Dostopno na: http://www.ena.com/oddelki/racunalninstvo/izd_3288_avt103875_streznik_hp_proliant_ml150g6_e5506
- [5] (2012) Google Drive. »Google Drive.« Dostopno na: <http://drive.google.com/start#home>
- [6] (2012) Infrastruktura kot storitev. Dostopno na: <http://infosrc.wordpress.com/2011/09/08/infrastruktura-kot-storitev-2/>
- [7] (2012) L.Kink, Zagotavljanje kakovosti podatkov v sistemih poslovne inteligence, magistrsko delo, Ljubljana, 2004. Dostopno na: <http://www.cek.ef.uni-lj.si/magister/kink325.pdf>
- [8] J.Merljak, *Računalništvo v oblaku*, diplomsko delo, Ljubljana, 2009. Dostopno na: http://eprints.fri.uni-lj.si/921/1/Merljak_J-UN.pdf
- [9] (2012) Mlacom. »NAS enota Thecus.« Dostopno na: <http://www.mlacom.si/thecus-n4100evo>

- [10] (2012) Mlacom. »Trdi disk WD.« Dostopno na:
<http://www.mlacom.si/trdi-diski-notranji/wd-wd2003fyys-2tb-7-2/64m/s300-raidedit>
- [11] (2012) Oblačno računalništvo v izobraževanju. Dostopno na:
<http://student.fnm.uni-mb.si/~zmocivnik/>
- [12] (2012) B.Osenar, Zagotavljanje kakovosti transakcijskih podatkov o kupcih za analitične potrebe, magistrsko delo, Ljubljana, 2010. Dostopno na:
<http://www.cek.ef.uni-lj.si/magister/osenar441-B.pdf>
- [13] (2012) Petrol. »Cenik elektrike in primerjava.« Dostopno na:
<http://www.petrol.si/energija-za-dom/energija/elektricna-energija/cenik-elektrike-primerjava>
- [14] (2012) Računalništvo v oblaku. Dostopno na:
http://www.zdruzenje-manager.si/storage/6766/12_Racunalnistvo_v_oblakih.pdf
- [15] (2012) Računalništvo v oblaku, fleksibilnejši dostop do računalniških storitev. Dostopno na:
<http://mladipodjetnik.si/podjetniski-koticek/poslovanje/racunalnistvo-v-oblaku-fleksibilnejši-dostop-do-racunalniskih-storitev>
- [16] (2012) Slika. »Tri glavna področja.« Dostopno na:
<http://www.13euro.com/cloud-computing.html>
- [17] (2012) SRC. Računalništvo v oblaku. Dostopno na:
http://www.src.si/library_si/pdf/infosrc/2012-68/racunalnistvo_v_oblaku.pdf
- [18] (2012) Wikipedia. »Cloud computing.« Dostopno na:
http://en.wikipedia.org/wiki/Cloud_computing
- [19] (2012) Wikipedija. »Platform as a service.« Dostopno na:
http://en.wikipedia.org/wiki/Platform_as_a_service
- [20] (2012) Wikipedia. »RAID.« Dostopno na:
<http://sl.wikipedia.org/wiki/RAID>
- [21] (2012) Wikipedija. »Software as a service.« Dostopno na:
http://en.wikipedia.org/wiki/Software_as_a_service

- [22] (2012) Wikipedija. »Transaction data.« Dostopno na:
http://en.wikipedia.org/wiki/Transaction_data
- [23] B.Brenčič, *Računalništvo v oblaku: stanje v Sloveniji in primerjava s tujino*, diplomsko delo, Ljubljana, 2010