

Analiza dostupnosti i modela kompenzacije u SLA ugovorima vodećih pružatelja *cloud* usluga

Autori:

Denis Subašić, Veleučilište u Rijeci

Sabrina Šuman, Veleučilište u Rijeci

Ivan Šimac, Veleučilište u Rijeci

Sažetak rada

Objasniti što je cloud računalstvo („računalstvo u oblaku”)

Objasniti modele pružanja usluga (IaaS, SaaS, PaaS)

Objasniti kvalitetu usluge u cloud sustavima (*engl. Quality of Services*) sa posebnim naglaskom na dostupnost

Analiza SLA (*engl. Service Level Agreement*) vodećih pružatelja usluga poput **Amazon Web Services**, **Microsoft Azure** i **Google Cloud Platform** s naglaskom na izračun dostupnosti i kompenzacijskih kredita

Usporedba SLA ugovora vodećih pružatelja usluga

Diskusija o obećanjima u praksi

Uvod – *cloud computing* ili računalinstvo u oblaku

Važna grana IT industrije koja ima široku primjenu u privatne i poslovne svrhe

Model isporuke računalnih resursa putem interneta (infrastruktura, platforma i aplikacije)

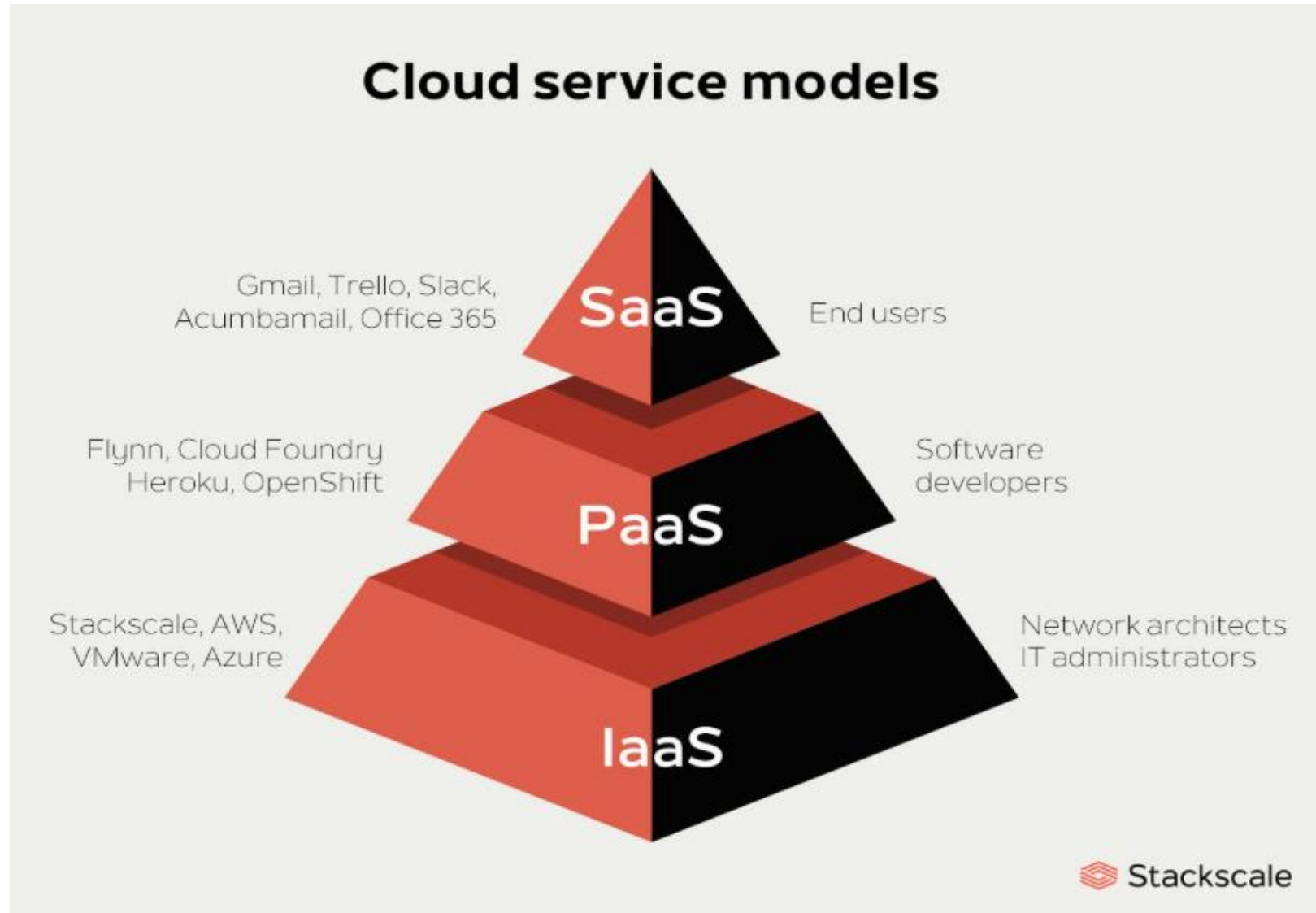
Podaci i aplikacije nalaze se u udaljenim podatkovnim centrima (serverima), a korisnik pristupa vlastitim računalom na poslužitelj

Omogućuje dinamičku dodjelu resursa prema potrebama korisnika

Prednosti: skalabilnost, fleksibilnost i manja opterećenost korisnikovih hardverskih resursa

Temeljni faktor kvalitete: **dostupnost** usluge

Modeli pružanja usluga



Izvor: <https://www.ir.com/guides/iaas-paas-saas>

Infrastructure as a Service (IaaS)

Najniža razina cloud modela

Pruža osnovne računalne resurse poput: virtualnih poslužitelja, mrežne infrastrukture, prostor za pohranu podataka

Pružatelj usluge upravlja: fizičkom infrastrukturom i virtualizacijom

Korisnik upravlja: operacijskim sustavom, aplikacijama i podacima

Često se koristi kod razvoja aplikacija i testnih okruženja

Primjeri: Amazon EC2, Microsoft Azure Virtual Machines, Google Compute Engine

Platform as a Service (PaaS)

Pružá razvojno okruženje za izradu aplikacija

Pružatelj upravlja infrastrukturom i operacijskim sustavom

Korisnik razvija i upravlja aplikacijom

Primjer: Google App Engine, AWS Elastic Beanstalk, Azure App Service

Software as a Service (SaaS)

Najviša razina apstrakcije u cloud modelima

Korisnici koriste gotove aplikacije putem interneta

Pružatelj upravlja infrastrukturom, platformom i aplikacijom

Korisnik koristi funkcionalnosti aplikacije bez upravljanja sustavom

Primjeri: Google Drive, Microsoft 365, Salesforce

Što se ide više prema SaaS modelu, odgovornost korisnika se smanjuje, dok pružatelj preuzima veći dio upravljanja sustavom.

Kvaliteta usluga u cloud sustavima

Kvaliteta cloud usluga ključna je za stabilan i siguran rad informacijskih sustava.

Cloud infrastruktura koristi se za pohranu i obradu osjetljivih podataka (osobni i financijski podaci).

Visoka razina kvalitete utječe na sigurnost podataka, poslovanje i korisničko iskustvo

Kvaliteta se najčešće procjenjuje kroz koncept Quality of Service (QoS), koji mjeri performanse sustava.

QoS parametri u cloud sustavima

Latencija - vrijeme odziva sustava

Propusnost - količina podataka prenesena u jedinici vremena

Pouzdanost - stabilan rad sustava bez grešaka

Skalabilnost - prilagodba sustava povećanom opterećenju

Performanse - učinkovitost izvršavanja zadataka

Sigurnost - zaštita podataka i sustava

Dostupnost - dostupnost usluge korisnicima

SLA ugovori u cloud sustavima

Service Level Agreement (SLA) definira odnos između pružatelja cloud usluge i korisnika te određuje razinu kvalitete usluge.

SLA ugovor uključuje minimalno zajamčene performanse sustava, odgovornosti pružatelja i korisnika i pravila u slučaju neispunjavanja usluge.

Ključni faktor SLA ugovora je definiranje dostupnosti usluge, te prag za kompenzaciju kredita.

Kompenzacijski krediti izraženi u dolarima (\$) za buduća plaćanja - nije stvarni povrat novčanih sredstava

Dostupnost kao ključni QoS parametar

- Dostupnost = udio vremena kada je usluga operativna
- I mala razlika u postotku znači puno minuta ili sati godišnje ili odabranom vremenskom periodu
- Pragovi u praksi: 99 % – 99,99 %
- Važno:

dostupnost usluge $\neq$ dostupnost aplikacije korisnika

Analiza SLA ugovora vodećih pružatelja usluga

- Izvori: službeni SLA dokumenti (AWS, Azure, Google Cloud)
- Fokus: IaaS (EC2, Azure VM, Compute Engine)
- Analiza:
 - *izračun dostupnosti,*
 - *pragovi, kompenzacije*
- Cilj: razumjeti što SLA realno pokriva

AWS SLA – Amazon EC2- Elastic Compute Cloud

- Amazon EC2: regionalni i instancijski SLA
- Regionalno: $\geq 99,99$ % (više zona u istoj regiji)
- Instanca: $\geq 99,5$ % mjesečno
- Kompenzacija: servisni kredit uz uvjete i izuzetke
- Napomena- SLA se ne primjenjuje na više sile, vanjske mrežne probleme ili greške korisnika

AWS SLA – modeli kompenzacije

Mjesečni postotak dostupnosti	Postotak kredita za uslugu
Manje od 99,99%, ali jednako ili veće od 99,0%	10%
Manje od 99,0%, ali jednako ili veće od 95,0%	30%
Manje od 95,0%	100%

Tablica 1. Kompenzacijske usluge AWS-a na razini regije, izvor: <https://aws.amazon.com/compute/sla/>

Postotak vremena rada na razini instance	Postotak kredita za uslugu
Manje od 99,5%, ali jednako ili veće od 99,0%	10%
Manje od 99,0%, ali jednako ili veće od 95,0%	30%
Manje od 95,0%	100%

Tablica 2. Kompenzacijske usluge AWS-a na razini instance, izvor: <https://aws.amazon.com/compute/sla/>

Azure Virtual Machines SLA

- Dostupnost ovisi o arhitekturi implementacije
- Availability Zone: $\geq 99,99$ % (najmanje 2 instance u različitim zonama)
- Availability Set: $\geq 99,95$ %
- Pojedinačna instanca: do 99,9 % (ovisno o disku)

Azure VM – modeli kompenzacije

Uptime Percentage	Service Credit
< 99.99%	10%
< 99%	25%
< 95%	100%

Tablica 3. Primjena kompenzacijskih kredita za Availability Zone,

Izvor: <https://www.microsoft.com/licensing/docs/view/Service-Level-Agreements-SLA-for-Online-Services?>

Uptime Percentage	Service Credit
< 99.95%	10%
< 99%	25%
< 95%	100%

Tablica 4. Primjena kompenzacijskih kredita za Availability Set,

Izvor: <https://www.microsoft.com/licensing/docs/view/Service-Level-Agreements-SLA-for-Online-Services?>

Uptime Percentage (Premium SSD, Premium SSD v2, and Ultra Disk)**	Uptime Percentage (Standard SSD Managed Disk)	Uptime Percentage (Standard HDD Managed Disk)	Service Credit
< 99.9%	< 99.5%	< 95%	10%
< 99%	< 95%	< 92%	25%
< 95%	< 90%	< 90%	100%

Tablica 5. Primjena kompenzacijskih kredita za pojedinu instancu

Izvor: <https://www.microsoft.com/licensing/docs/view/Service-Level-Agreements-SLA-for-Online-Services?>

SLA za Google Compute Engine

- Dvije mrežne razine: Premium i Standard
- Više zona: do $\geq 99,99$ % (ovisno o razini i tipu instance)
- Pojedinačne instance: tipično $\geq 99,9$ %
- Za kredit: zahtjev podršci i dokaz u roku (npr. 60 dana)

GCP – financijski kredit (više zona / load balancing)

Mjesečni postotak dostupnosti	Postotak mjesečnog računa za odgovarajuću Pokrivenu uslugu u Regiji koja nije ispunila SLO, a koji će biti uračunat u buduće mjesečne račune Kupca
99,00% - < 99,99% (Premium razina) 99,00% - < 99,9% (Standardna razina)	10%
95,00% - < 99,00% (Premium i Standardna razina)	25%
< 95,00% (Premium i Standardna razina)	100%

Tablica 6. Krediti za više zona i load balancing (izuzete pojedine regije)
Izvor: <https://cloud.google.com/compute/sla>

GCP – financijski kredit (memory-optimized)

- Tablica: krediti za memory-optimized instance (izuzete pojedine regije)

Mjesečni postotak dostupnosti	Postotak mjesečnog računa za jednu instancu u regiji koja nije ispunila SLO, a koji će biti uračunat u buduće mjesečne račune Kupca
95,00% - < 99,95%;	10%
90,00% - < 95,00%	25%
< 90,00%	100%

Tablica 6. krediti za memory-optimized instance (izuzete pojedine regije)

Izvor: <https://cloud.google.com/compute/sla>

Usporedba karakteristika vodećih pružatelja usluga

Karakteristika	AWS EC2	Azure Virtual Machines	Google Compute Engine
Organizacija infrastrukture	Regije i Availability Zone (AZ)	Regije, Availability Zone i Availability Set	Regije i Zone
SLA za visoku dostupnost	99,99 % za instance raspoređene u najmanje dvije AZ	99,99 % za instance u dvije Availability Zone	99,99 % za instance u više zona
SLA za srednju razinu redundancije	Nije posebno definiran	99,95 % za Availability Set	Nije posebno definiran
SLA za pojedinačnu instancu	99,50%	do 99,9 % (ovisno o vrsti diska)	99,90%
Definicija nedostupnosti	Gubitak vanjske mrežne povezanosti instance	Gubitak Virtual Machine Connectivity (TCP/UDP komunikacija)	Gubitak vanjske povezivosti ili pristupa disku
Izračun dostupnosti	100 % – postotak downtime minuta	(Ukupne minute – downtime) / ukupne minute × 100	(Ukupne minute – downtime) / ukupne minute × 100
Kompenzacija	Service credits	Service credits	Financijski krediti
Raspon kompenzacije	10 % – 100 %	10 % – 100 %	10 % – 100 %
Način isplate	Kredit za buduće korištenje usluge	Kredit za buduće korištenje usluge	Kredit za buduće račune
Zahtjev za kompenzaciju	Korisnik mora podnijeti zahtjev AWS-u	Korisnik mora podnijeti zahtjev Microsoftu	Korisnik mora kontaktirati Google podršku (rok 60 dana)
SLA izuzeća	Događaji izvan kontrole AWS-a, korisničke greške	Korisnička konfiguracija i problemi izvan Azure infrastrukture	Faktori izvan Google kontrole, softver korisnika ili treće strane

Izvor: službeni SLA ugovori AWS, Azure i GCP uz obradu autora

Kolika su obećanja u praksi?

- SLA ugovori definiraju očekivanu razinu dostupnosti usluga, ali ne jamče potpuno neprekidan rad sustava. - SLA $\neq$ jamstvo neprekidnog rada
- Veliki cloud pružatelji koriste distribuirane podatkovne centre uz jaku infrastrukturu kako bi smanjili rizik prekida rada.
- Unatoč tome, tehnički kvarovi, cyber napadi i ljudske pogreške i dalje predstavljaju moguće uzroke nedostupnosti usluga.
- Zbog toga SLA treba promatrati prvenstveno kao ugovorni mehanizam za definiranje odgovornosti i kompenzacije, a ne kao apsolutno jamstvo pouzdanosti sustava.
- Za realnu otpornost: arhitektura, monitoring i plan oporavka

Zaključak

- Cloud računalstvo danas je standard u poslovanju, a modeli usluga (IaaS/PaaS/SaaS) određuju podjelu odgovornosti između korisnika i pružatelja te se kvaliteta prati kroz ključne QoS pokazatelje.
- Analiza SLA ugovora za AWS EC2, Azure Virtual Machines i Google Compute Engine pokazuje visoku dostupnost do 99,99%, uz razlike u pragovima i uvjetima koje najčešće ovise o višezonskoj implementaciji i dodatnim resursima.
- Najviši pragovi zahtijevaju višezonsku implementaciju
- Kršenje SLA-a → krediti, uz izuzetke i proceduru
- SLA je okvir odgovornosti, ne apsolutno jamstvo

AUTORI

Denis Subašić, bacc.inf.

student stručnog diplomskog studija Informacijske tehnologije u poslovnim sustavima

Odjel za Informacijske i komunikacijske tehnologije

Veleučilište u Rijeci

dsubasic@veleri.hr

doc. dr. sc. Sabrina Šuman, prof.struč.stud.

Odjel za Informacijske i komunikacijske tehnologije

Veleučilište u Rijeci

ssuman@veleri.hr

Ivan Šimac, pred.

Odjel za Informacijske i komunikacijske tehnologije

Veleučilište u Rijeci

isimac@veleri.hr



**HVALA NA
PAŽNJI!**



PITANJA?