Univerzitet u Nišu,

Elektronski fakultet Niš

Logo

Description automatically generated

*Seminarski rad*

Sistemi za upravljanje bazama podataka

**CLOUD BAZE PODATAKA I DATABASE-AS-A-SERVICE REŠENJA – CLOUD SQL**

Mentor: Student:

Aleksandar Stanimirović Andrija Milosavljević 1481

Niš, jun 2023. godine

**Sadržaj**

[**1.** **Uvod** 3](#_Toc137313913)

[**2.** **Cloud SQL** 5](#_Toc137313914)

[**2.1.** **Karakteristike Cloud SQL-a** 6](#_Toc137313915)

[**2.2.** **SQL Server na Cloud SQL-u** 7](#_Toc137313916)

[**3.** **Zaključak** 14](#_Toc137313917)

[**4.** **Literatura** 15](#_Toc137313918)

# **Uvod**

Cloud baze podataka su vrsta baza podataka koje se smeštaju i upravljaju u cloud okruženju, umesto na lokalnim serverima. One se često koriste za skladištenje, upravljanje i obradu velikih količina podataka. Cloud baze podataka pružaju elastičnost u pogledu kapaciteta skladištenja i obrade podataka. Lako se mogu povećati ili smanjiti resursi u skladu sa potrebama, što omogućava prilagođavanje promenljivim zahtevima i optimizaciju troškova.

Cloud baze podataka često podržavaju funkcionalnost replikacije podataka, što omogućava bolju dostupnost podataka i otpornost na kvarove. One pružaju različite modele baza podataka, uključujući relacione baze podataka, NoSQL baze podataka ili graf baze podataka.

Postoje dva modela okruženja cloud baze podataka, tradicionalni model oblaka i baza podataka kao usluga (eng. *Database-as-a-service* – *DBaaS*). U tradicionalnom modelu oblaka, organizacija ili korisnik koristi cloud infrastrukturu za implementaciju i upravljanje sopstvenim bazama podataka. Organizacija ima veću kontrolu nad infrastrukturom i instalira i konfiguriše bazu podataka prema sopstvenim potrebama. Cloud provajder obezbeđuje virtuelne mašine, skladište i mrežu za podršku ovim bazama podataka, ali administracija i održavanje baze podataka uglavnom su odgovornost korisnika [1].

Sa druge strane, *Database-as-a-service* je usluga računarstva u oblaku koja omogućava korisnicima pristup i korišćenje baze podataka u oblaku, bez kupovine, instalacije i postavljanja baze podataka na svoju infrastrukturu. *DBaaS* radi na infrastrukturi provajdera usluga i oni su odgovorni za sve probleme ili greške koje bi se mogle dogoditi. Korisnik ne mora brinuti o instalaciji, konfiguraciji i održavanju baze podataka, već te odgovornosti preuzima provajder. Korisnik može jednostavno koristiti API za upravljanje i pristup podacima.

*DBaaS* pruža mnoge prednosti kao što su [1]:

* Ušteda troškova – postavljanje infrastrukture za upravljanje bazom podataka je skupo, dok je korišćenjem *DBaaS* usluge potrebno plaćati periodičnu naknadu na osnovu potrošenih resursa
* Skalabilnost – prilikom povećane upotrebe aplikacije, dostupno je automatsko skaliranje instance baze podataka. Na primer, ukoliko upotreba poraste preko određenog praga, podaci sa jedne instance se mogu automatski distribuirati na više instanci
* Jednostavno upravljanje – za upravljanje i održavanje baze podataka koje su on-premise je potrebno imati tim ljudi koji se bavi održavanjem, dok se u oblaku provajder brine o upravljanju baze podataka, uključujući sigurnost, rezervne kopije, nadogradnje, odražavanje sistema
* Brži razvoj i kraće vreme do tržišta – *DBaaS* usluga poboljšava vreme obezbeđivanja baze podataka automatizacijom procesa. Jednom kada se obezbedi i konfiguriše baza podataka, ona je spremna da se integriše sa aplikacijom kroz nekoliko minuta
* Bezbednost podataka – provajderi baza podataka u oblaku pružaju sigurnosne funkcionalnosti kao što su enkripcija podataka, virtuelne privatne mreže i sigurnosne grupe

Međutim, važno je napomenuti da *DBaaS* ima i neke izazove, kao što su bezbednost i privatnost podataka. Korisnici moraju pažljivo odabrati pouzdanog provajdera sa dobrim bezbedonosnim praksama i mehanizmima zaštite podataka.

Uprkos izazovima, *DBaaS* se često koriste kao praktično i ekonomično rešenje za mnoge organizacije. Prema istraživanju kompanije Gartner iz 2022. godine, procenjuje se da više od 75% organizacija koristi cloud baze podataka, što ukazuje na veliki rast ove industrije. Neki od najpoznatijih cloud provajdera, a ujedno i vodeći u ovoj industriji su: Amazon Web Services (AWS), Microsoft Azure, Google Cloud i Oracle [2].

Što se tiče tržišta, trenutno je najdominantnija Severna Amerika sa udelom od 35,1% u 2022. godini i očekuje se da će zadržati svoju dominaciju u narednom periodu (*Slika 1.1.*). Glavni faktor koji utiče na rast tržišta je sve veće usvajanje cloud tehnologija u raznim industrijama. Organizacije sve više prepoznaju prednosti računarstva u oblaku, što ih podstiče da migriraju podatke na cloud platforme. Sa rastom podataka, postoji sve veća potreba za pouzdanim i skalabilnim rešenjima za skladištenje i obradu podataka [3].

A picture containing text, screenshot, map

Description automatically generated

Slika 1.1. Tržište cloud baza podataka i DBaaS

# **Cloud SQL**

Cloud SQL je potpuno upravljana usluga za relacione baze podataka koju pruža Google Cloud. Ona omogućava korisnicima da jednostavno koriste MySQL, PostgreSQL i SQL Server baze podataka u cloud okruženju. Cloud SQL preuzima zadatke administracije baze podataka, uključujući instalaciju, konfiguraciju, sigurnost, skaliranje, rezervne kopije. Ovo omogućava korisnicima da se fokusiraju na upravljanje svojim podacima umesto na zadacima administracije.

Cloud SQL pruža visoku dostupnost i skalabilnost tako što obezbeđuje automatsku replikaciju podataka i podršku za oporavak u slučaju kvarova, a u situacijama kada se javi potreba za dodatnim resursima, moguće je automatsko skaliranje. Takođe, on se lako integriše sa drugim uslugama koje pruža Google Cloud, kao što su Google Kubernetes Engine (GKE), Google App Engine (GAE), BigQuery i dr.

Jedan od ključnih termina u Cloud SQL-u je Cloud SQL instanca. Svaka instanca se pokreće na virtuelnoj mašini koja radi na serveru Google Cloud-a. Svaka virtuelna mašina pokreće bazu podataka (MySQL/PostgreSQL/SQL Server), i servise koji pružaju podršku, kao što su beleženje aktivnosti (eng. Logging), praćenje performansi (eng. Monitoring), itd. Baza podataka se skladišti na skalabilnom i trajnom disku (eng. PersistentDisk) koji se povezuje sa virtuelnom mašinom. Statička IP adresa se dodeljuje svakoj virtuelnoj mašini kako bi se osiguralo da IP adresa sa kojom se aplikacija povezuje ostaje ista tokom celog životnog ciklusa Cloud SQL instance (*Slika 2.1.*) [4].

A screenshot of a computer

Description automatically generated

Slika 2.1. Arhitektura Cloud SQL instance

## **Karakteristike Cloud SQL-a**

**Rezervne kopije** – omogućavaju vraćanje izgubljenih/oštećenih podataka u Cloud SQL instancu. Pored toga, ukoliko na instanci dođe do greške, moguće je vratiti na prethodno stanje korišćenjem kreiranih rezervnih kopija. Za potrebe kreiranja rezervnih kopija, za svaku instancu se kreira novi korisnik *cloudsqladmin*.

Postoje dve vrste rezervnih kopija: na zahtev i automatske. Rezervne kopije na zahtev (eng. On-demand backup) se pokreću ručno i mogu se kreirati u bilo kom trenutku. Kada se pokrene kreiranje kopije, Cloud SQL pravi kopiju trenutnog stanja baze podataka. One su korisne u situacijama kada korisnik želi da napravi kopiju podataka pre nego što izvrši neke važne promene ili ažuriranja na bazi podataka.

Automatske rezervne kopije (eng. Automated backup) se kreiraju svakodnevno u vremenskom periodu koji traje 4 časa. Preporuka je da se kreiranje kopija zakazuje kada baza podataka ima najmanje aktivnosti. Maksimalan broj automatskih rezervnih kopija je konfigurabilan, od 1 do 365. Cloud SQL pravi rezervnu kopiju instance samo ako su se podaci instance promenili od poslednje rezervne kopije.

**Replikacija** – mehanizam kopiranja podataka u više instanci čime se postiže redundantnost podataka. Cloud SQL ima klaster okruženje za replikaciju koje se sastoji od primarne instance, rezervne instance, i jedne ili više replika za čitanje. Primarni razlog za korišćenje replikacije je skaliranje podataka u bazi bez pogoršanja performansi. Drugi razlozi uključuju migraciju podataka između regiona, kao i migraciju iz neke privatne baze podataka na Cloud SQL. Replike za čitanje se koriste kako bi se rasteretila primarna instanca. Sve promene na primarnoj instanci se ažuriraju u realnom vremenu na replikama za čitanje.

**Visoka dostupnost** (eng. High availability) – kada se konfiguriše Cloud SQL instanca sa visokom dostupnošću, sistem automatski kreira rezervnu instancu u drugoj zoni koja je identična primarnoj instanci. U slučaju kvara, rezervna instanca postaje nova primarna instanca. Korisnici se tada preusmeravaju na novu primarnu instancu. Ovaj proces se naziva preusmeravanje (eng. failover) i obavlja se automatski kako bi obezbedio minimalno vreme nedostupnosti podataka. Dakle, cilj je smanjiti vreme nedostupnosti kada zona ili instanca postanu nedostupni. To se može dogoditi tokom prekida u zoni ili kada instance ostane bez memorije.

**Uvoz/izvoz podataka** – omogućava uvoz podataka iz SQL datoteka koja sadrži naredbe za kreiranje tabele ili unos podataka. Ovo je pogodno za prenos već pripremljenih SQL skripti u Cloud SQL. Takođe, moguć je uvoz podataka iz drugih izvora koje pruža Google, kao što su Google Cloud Storage, BigQuey ili neka druga Cloud SQL instanca. Što se tiče izvoza podataka, moguće je izvesti sadržaj baze podataka u SQL datoteci koja se može koristiti kao sigurnosna kopija ili za prenos podataka na drugu lokaciju [4].

## **SQL Server na Cloud SQL-u**

Cloud SQL-om je moguće upravljati interfejsno pomoću Admin API-ja ili pomoću *gcloud* komandne linije. Koristeći Admin API možete kreirati, konfigurisati, upravljati i nadgledati Cloud SQL instance i resurse. API pruža funkcionalnosti kao što su kreiranje, brisanje, ažuriranje instanci, upravljanje korisnicima i privilegijama, pravljenje sigurnosnih kopija, upravljanje replikama itd. Sa druge strane, *gcloud* je komandna linija koja pruža upravljanje i administraciju putem terminala. Pomoću *gcloud* komandne linije možemo izvršiti iste funkcionalnosti kao i preko API-ja [6].

Na početku je neophodno kreirati Cloud SQL instancu. To se može izvršiti na dva načina, pomoću Admin API-ja (*Slika 2.2.*) ili *gcloud* komandne linije (*Slika 2.3.*). Prilikom kreiranja instance preko API-ja, neophodno je uneti jedinstveni identifikator instance, lozinku, region u kom želimo da se nalazi instanca i verziju SQL Servera. Pored ovih stvari, moguće je konfigurisati dodatne funkcionalnosti kao što su veličina skladišta, način održavanja, automatsko kreiranje rezervnih kopija, automatsko skaliranje, visoka dostupnost, itd. [5].

A screenshot of a computer

Description automatically generated

Slika 2.2. Kreiranje Cloud SQL instance pomoću Admin API-ja

A screenshot of a computer

Description automatically generated with medium confidence

Slika 2.3. Kreiranje Cloud SQL instance preko gcloud komandne linije

Pregled svih instanci možemo videti pomoću komande *gcloud sql instances list*, pri čemu ćemo dobiti informacije o nazivu instance, verziji baze podataka, lokaciji na kojoj se nalazi instanca, IP adresa instance, kao i status (*Slika 2.4.*).

A screenshot of a computer

Description automatically generated with medium confidence

Slika 2.4. Pregled svih kreiranih instanci

Nakon kreiranja Cloud SQL instance, možete se povezati na nju kako biste pristupili i upravljali bazom podataka. Povezivanje omogućava izvršavanje SQL upita, ažuriranje podataka, konfigurisanje sigurnosnih parametara i ostale administrativne zadatke. Postoje 2 načina povezivanja na Cloud SQL instancu:

* Povezivanje pomoću privatne IP adrese – omogućava aplikacijama da se direktno povežu sa instancom pomoću privatne IP adrese, koja je dostupna u virtuelnom privatnom oblaku (VPC), bez potrebe za korišćenjem javne ili prolaska preko interneta. Ovo pruža dodatnu sigurnost i smanjuje latenciju veze
* Povezivanje pomoću javne IP adrese – ovo je najčešći način povezivanja, jer omogućava korišćenje standardnih SQL klijenata za povezivanje na instancu, kao što je *Microsoft SQL Server Management Studio*. Kada se koristi javna adresa, postoji potreba za sigurnosnim mehanizmima koji kontrolišu pristup instanci i održavaju bezbednost podataka. Cloud SQL Auth Proxy je alat koji obezbeđuje sigurno povezivanje putem javne IP adrese, omogućavajući autentifikaciju korisnika i kontrolu pristupa

Da bi se korisnik povezao na instancu pomoću *Microsoft SQL Server Management Studio*-a, neophodno je da za parametar tip servera izabere opciju *Database Engine*, dok se za naziv servera koristi javna IP adresa instance. Za način autentifikacije se koristi opcija *SQL Server Authentication*, gde je potrebno navesti kredencijale koje smo unosili prilikom kreiranja instance (*Slika 2.5.*).

A screenshot of a computer

Description automatically generated

Slika 2.5. Povezivanje na instancu pomoću Microsoft SQL Management Studio-a

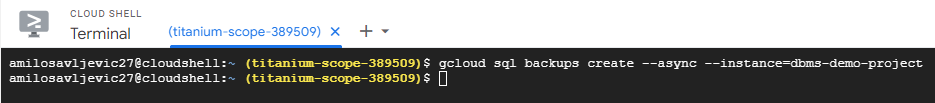
Nakon uspešnog povezivanja na instancu, možemo kreirati bazu podataka, tabele, unositi podatke pomoću T-SQL naredbi u *MS SQL Management Studio*-u. Drugi način za kreiranje baze je preko *gcloud* komandne linije ili interfejsno preko API-ja (*Slika 2.6.*). Svaki kreirani objekat, bilo iz *MSSQL Management Studio*-a ili preko komandne linije, će se automatski sinhronizovati u realnom vremenu, pa ćemo tako imati konzistentnu situaciju na Cloud SQL-u.

A screenshot of a computer

Description automatically generated

Slika 2.6. Kreiranje baze podataka pomoću API-ja i komandne linije

Kao što smo već rekli, postoje 2 vrste rezervnih kopija, automatske i na zahtev. Kreiranje kopija se može izvršiti preko komandne linije (*Slika 2.7.* i *2.8.*). Sve kreirane rezervne kopije je moguće pregledati preko API-ja ili ih izlistati komandom *gcloud sql backups list*, dok za pregled informacija o konkretnoj kopiji koristimo komandu *gloud sql backups describe*. Na *slici 2.9.* je prikazan primer upotrebe ovih komandi.



Slika 2.7. Kreiranje rezervne kopije na zahtev

A picture containing text, screenshot, font, information

Description automatically generated

Slika 2.8. Kreiranje automatske rezervne kopije

A screenshot of a computer

Description automatically generated

Slika 2.9. Pregled kreiranih rezervnih kopija iz API-ja i komandne linije

Replike u Cloud SQL-u predstavljaju kopije primarne instance koje se koriste za čitanje podataka, kako bi se smanjilo opterećenje. To omogućava bolje performanse za operacije čitanja, posebno u slučajevima kada je primarna instanca zauzeta upisivanjem podataka, a samim tim ovo povećava dostupnost podataka. Svaka replika se sinhronizuje sa bazom podataka kako bi se osigurala konzistentnost podataka.

Prilikom kreiranja replike za čitanje, instanca čija se replikacija vrši mora da bude primarna. U Cloud SQL-u postoji ograničenje za SQL Server baze podataka, pa je moguće kreirati repliku za čitanje samo nad *SQL Server 2017 Enterprise* i *SQL Server 2019 Enterprise*. Za jednu instancu je moguće kreirati najviše 8 replika za čitanje. Komanda za kreiranje replike je identična kao i za kreiranje instance, s tim što se navodi master (primarna) instanca na koju se replika odnosi (*Slika 2.10.*). Dodatne informacije o replikama možemo videti komandom *cloud sql instances describe*.

A screenshot of a computer

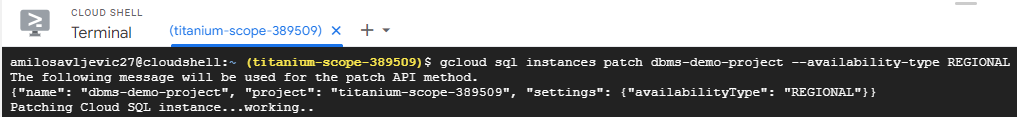
Description automatically generated with medium confidence

Slika 2.10. Kreiranje replike za čitanje i pregled

Prilikom kreiranja/ažuriranja instance postoji opcija za omogućavanje visoke dostupnosti (HA) instance, pri čemu se obezbeđuje da baza podataka bude dostupna i funkcionalna čak i u slučaju problema sa hardverom, softverom ili mrežom.

Omogućavanje HA instanci baze podataka se postiže kroz konfiguraciju regionalne instance (*Slika 2.11.*). Regionalna instanca je sastavljena od primarne instance i rezervne instance smeštene u različitim zonama unutar istog regiona. Primarna instanca je aktivna instanca koja obrađuje zahteve za upisivanje i čitanje podataka. Rezervna instanca je pasivna instanca koja se stalno sinhronizuje sa primarnom instancom i može preuzeti ulogu primarne instance u slučaju da primarna instanca postane nedostupna, a može se izvršiti i kasnije nad postojećom instancom.

Iz komandne linije, nakon podešavanja HA konfiguracije, može se inicirati preusmeravanje (eng. failover) da bi se proverilo kako će aplikacija da reaguje ukoliko dođe do toga (*Slika 2.12*). Dok se vrši preusmeravanje, primarna instanca nije dostupna i nije moguće povezati se na nju dok se preusmeravanje ne izvrši. Ako je preusmeravanje uspešno, nakon njegovog iniciranja korisnik može da se poveže i nastavi da koristi instancu.



Slika 2.11. Omogućavanje HA konfiguracije

A screenshot of a computer

Description automatically generated with medium confidence

Slika 2.12. Iniciranje preusmeravanja instance

Cloud Logging je usluga koja omogućava skladištenje, pretragu i analizu log podataka i podataka o događajima sa Google Cloud-a [8]. U Cloud SQL-u ova usluga omogućava praćenje i nadgledanje aktivnosti i događaja nad kreiranim instancama kako bi otkrili probleme, pratili performanse i povećali sigurnost. Cloud Logging takođe omogućava postavljanje upozorenja na osnovu određenih kriterijuma kako bi bili obavešteni o važnim događajima ili greškama u Cloud SQL instancama. Za pregled logova možemo koristiti komandu *gcloud logging read* ili Log Explorer koji pruža interaktivan pregled logova i manipulaciju nad log podacima (*Slika 2.13.*).

A screenshot of a computer

Description automatically generated with medium confidence

Slika 2.13. Pregled log podataka u Log Explorer-u

Cloud Monitoring omogućava nadgledanje i praćenje stanja, performansi i resursa Cloud SQL instanci. Pomoću kontrolnih tabli možemo pratiti ključne performanse kao što su opterećenje CPU-a, dostupnost memorije, broj konekcija i druge informacije o stanju instanci [7]. Pored toga, moguće je kreirati i sopstvene kontrolne table koje će prikazivati podatke koji su vama važni. Na *slici 2.14.* možemo videti kontrolnu tablu opterećenja CPU-a, pri čemu vidimo da instanca nije bila dostupna određeno vreme, tj. za vreme preusmeravanja iz primera na *slici 2.12*. Na drugom primeru možemo videti kontrolnu tablu korišćenja memorije u izabranom periodu (*Slika 2.15.*).

Korišćenjem ove funkcionalnosti možemo pratiti stanje i performanse baze podataka u realnom vremenu. Ove informacije nam mogu pomoći da identifikujemo potencijalne probleme i optimizujemo performanse.

A screenshot of a computer

Description automatically generated with low confidence

Slika 2.14. Praćenje opterećenje CPU-a

A red line on a white background

Description automatically generated with low confidence

Slika 2.15. Praćenje korišćenja memorije

# **Zaključak**

U današnjem svetu sve veća upotreba oblaka donosi sa sobom brojne prednosti, posebno u području upravljanja bazama podataka. Cloud baze podataka pružaju organizacijama skalabilnost, fleksibilnost i pouzdanost u upravljanju podacima. Uvođenje cloud baza podataka omogućava organizacijama da se oslobode složenih zadataka administracije, kao i da smanje troškove i fokusiraju se na upravljanje podacima.

Međutim, efikasno upravljanje i održavanje baze podataka može biti izazovno. Upravo se tu kao rešenje javlja baza podataka kao usluga (*DBaaS*) koja omogućava korisnicima pristup i upravljanje bazom podataka bez potrebe za sopstvenom infrastrukturom i administracijom. Provajder cloud usluge preuzima odgovornost za sigurnost, rezervne kopije i dostupnost baze podataka, oslobađajući korisnike od tih zadataka.

Jedan od predstavnika *DBaaS* je Google Cloud SQL koji pomaže korisnicima u upravljanju relacionih baza podataka, uključujući MySQL, PostgreSQL i SQL Server. Cloud SQL nudi brojne prednosti, uključujući automatsko skaliranje, replikaciju, visoku dostupnost i sigurnost. Korisnici mogu jednostavno povećati ili smanjiti kapacitet baze podataka prema svojim potrebama, bez zastoja u radu. Takođe mogu kreirati replike svojih baza podataka radi poboljšanja dostupnosti i performansi. Omogućeno je i automatsko kreiranje rezervnih kopija koje se koriste za vraćanje izgubljenih/oštećenih podataka ukoliko dođe do neke greške ili kvara. Sve ovo omogućava organizacijama da efikasno upravljaju bazama podataka (koje su u konstantnom rastu) i obezbeđuje neprekidan pristup i dostupnost podacima bilo gde u svetu.

# **Literatura**

1. What is a cloud database? An in-depth cloud DBMS guide, <https://www.techtarget.com/searchcloudcomputing/definition/cloud-database>, pristupljeno: 3.6.2023.
2. Cloud Databases Are Maturing Rapidly, Gartner Says, <https://www.datanami.com/2023/01/23/cloud-databases-are-maturing-rapidly-gartner-says/>, pristupljeno: 3.6.2023.
3. Cloud Database And DBaaS Market Size, <https://www.grandviewresearch.com/industry-analysis/cloud-database-dbaas-market-report>, pristupljeno: 4.6.2023.
4. Cloud SQL Documentation, <https://cloud.google.com/sql/docs/sqlserver/introduction>, pristupljeno: 6.5.2023.
5. SQL Server on Google Cloud, <https://cloud.google.com/sql-server>, pristupljeno: 11.6.2023.
6. Cloud CLI: gcloud, <https://cloud.google.com/sdk/gcloud/reference>, pristupljeno: 11.6.2023.
7. Cloud Monitoring, <https://cloud.google.com/monitoring/docs/monitoring-overview>, pristupljeno: 18.6.2023.
8. Cloud Logging, <https://cloud.google.com/logging/docs/overview>, pristupljeno: 18.6.2023.