UNIVERZITET U NISU

ELEKTRONSKI FAKULTET

**IZOLACIJA I KONKURENTNI PRISTUP KOD POSTGRESQL BAZE PODATAKA**

Seminarski rad

Studijski program: Racunarstvo i informatika

Modul: Softversko inzenjerstvo

Mentor: Student:

Prof. dr Aleksandar Stanimirovic Zeljko Vasic, br.ind. 1808

Sadrzaj

[1. Uvod 3](#_Toc4007)

[2. Nivoi izolacije u bazama podataka 4](#_Toc20769)

[2.1. Read uncommitted 4](#_Toc7472)

[2.2. Read commited 4](#_Toc27263)

[2.3. Repeatable read 6](#_Toc1359)

[2.4. Serializable 9](#_Toc10010)

[3. Primeri anomalija zbog kojih je izolacija potrebna 9](#_Toc15285)

[4. Mehanizmi za obezbedjivanje konkurentnog pristupa 13](#_Toc24034)

[4.1. MVCC 13](#_Toc11659)

[4.2. Snapshot mehanizam 13](#_Toc22895)

[4.3. Row locking 13](#_Toc3290)

[4.4. Table locking 16](#_Toc30280)

[4.5. Predicate locking 16](#_Toc29727)

[5. Zakljucavanje i upravljanje konkurentnim pristupom 17](#_Toc15525)

[5.1. Pessimistic locking 17](#_Toc2976)

[5.2. Optimistic locking 19](#_Toc6395)

[5.3. Problemi u upravljaju zakljucavanjem 22](#_Toc26167)

[5.4. Zakljucavanje i performanse 24](#_Toc12809)

[6. Zakljucak 25](#_Toc14770)

[7. Literatura 26](#_Toc4169)

# Uvod

Izolacija obezbedjuje da transakcije koje se izvrsavaju istovremeno uticu jedna na drugu na predvidiv nacin. Fokusira se na to kako ce se transakcije ponasati kada istovremeno pristupaju istim resursima. Izolacija odredjuje koliko su transakcije medjusobno izolovane, odnosno odvojene jedne od drugih i u kojoj meri im je dozvoljeno da vide promene drugih transakcija koje se trenutno izvrsavaju.

Vecina danasnjih baza podataka omogucava konkurentno izvrsavanje transakcija nad bazom. To sa sobom donosi rizik da se podaci promene izmedju dva citanja u okviru iste transakcije, da se citaju nepotvrdjene vrednosti iz drugih transakcija ili da se podaci menjaju istovremeno od strane vise transakcija. Izolacija omogucava kontrolisanje anomalija koje ugrozavaju konzistentnost baze podataka.

Najcesce anomalije su:

* Dirty read - nastaje kada jedna transakcija cita necommitovane podatke druge transakcije, odnosno podatke koji jos uvek nisu potvrdjeni i potencijalno mogu biti ponisteni.
* Non-repetable read - nastaje kada jedna transakcija izvrsava isti upit za citanje vise puta, ali dobija razlicite rezultate. To se desava zato sto je druga transakcija, dok se prva jos uvek izvrsavala, izmenila te podatke u bazi i prva transakcija dobija drugaciji rezultat nego prilikom prvog citanja.
* Lost update - nastaje kada dve transakcije menjaju istovremeno isti red, i jedna transakcija pregazi promene druge transakcije.
* Phantom read - nastaje slicno kao i Non-repetable read osim sto druga transakcija u medjuvremenu dodaje redove koji odgovaraju filterima upita za citanje, i prva transakcija dobija vise redova u drugom citanju nego u prvom citanju.
* Write skew - nastaje kada dve transakcije istovremeno menjaju razlicite redove i nastaje logicki problem. Nema direktnog fizickog konflikta ali kada se spoje promene te dve transakcije narusava se integritet podataka.

Izolacija je jedno od cetiri osnovna svojstva koje cine poznati ACID model(Atomocity, Consistency, Isolation, Durability).

Atomocity(Atomicnost) je svojstvo koje osigurava da ce se transakcija izvrsiti u celini ili se nece izvrsiti uopste. Transakcija se gleda kao celina, i ne moze se deliti, odnosno nije moguce izvrsiti samo jedan deo transakcije.

Consistency(Konzistentnost) je svojstvo da se baza prevodi iz jednog konzistentnog stanja u drugo konzistentno stanje. Odnosno baza nikada ne moze da se nadje u nekonzistentnom stanju. Nekonzistentna baza je ona baza ciji su podaci nepotpuni, kontradiktorni ili nevazeci.

Durability(Trajnost) je svojstvo da promene u bazi koje su jednom potvrdjene ostaju trajno zabelezene u bazi, ne gube se cak ni ako baza padne.

# Nivoi izolacije u bazama podataka

Nivo izolacije odredjuje koliko su transakcije izolovane jedna od druge, odnosno koje anomalije konkurentnog pristupa su dozvoljene na tom nivou izolacije a koje nisu.

SQL standard definise 4 standardna nivoa izolacije:

1. Read uncommitted
2. Read committed
3. Repetable Read
4. Serializable

## Read uncommitted

Najslabiji nivo izolacije. Svaka od nabrojanih anomalija se moze desiti na ovom nivou. Transakcije mogu da vide nepotvrdjene(unccommited) promene drugih transkacije. To znaci da se moze desiti dirty read, odnosno da jedna transakcija cita podatke koje je druga transakcija izmenila ali nije commitovala i moze da izvrsi rollback. Takodje je moguce da podaci koje koristi jedna transkacija budu izmenjeni od strane druge transkacije, u toku trajanja prve transakcije, i tada dolazi do nonrepetable reads (isti red se menja izmedju dva citanja) ili phantom reads (novi redovi se pojave izmedju dva citanja sa istim kriterijumom).

PostgreSQL ne implementira ovaj nivo.

## Read commited

Read commited je podrazumevani nivo izolacije u vecini baza podataka (ukljucujuci i PostgreSQL). Svaki SQL upit vidi samo potvrdjene podatke, odnosno upit u trenutnoj transakciji ne moze da procita podatke koji su izmenjeni od strane druge transakcije a jos uvek nisu commitovani (dirty read je sprecen na ovom nivou izolacije). Ipak, i dalje se desava problem da ako se u jednoj transakciji jedan upit izvrsi vise puta rezultat se moze promeniti ako neka druga transkacija u medjuvremenu upise i potvrdi nove vrednosti.

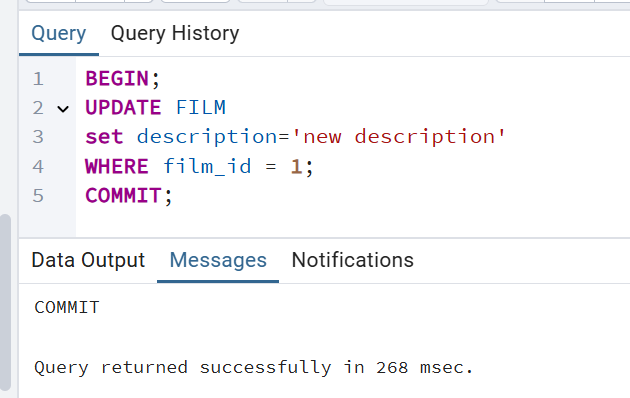
Dozvoljene anomalije su i dalje non-repetable reads, lost-update, phantom reads i write-skew.

Non-repetable read primer:

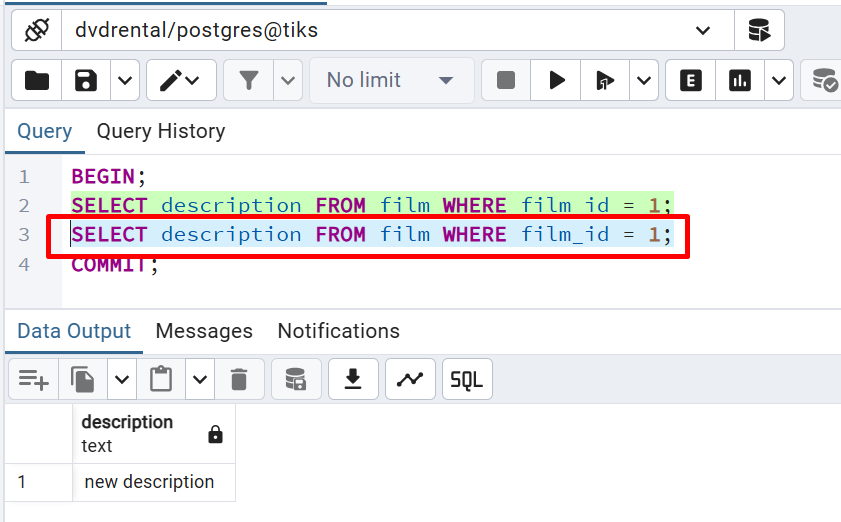
Prva transakcija pocela sa izvrsavanjem, i izvrsio se samo oznaceni deo, vidimo da je upit vratio odredjeni podatak:



U medjuvremenu druga transakcija se izvrsava u celosti i menja bas taj podatak koji prva transakcija cita:



Prva transakcija nastavlja sa izvrsavanjem, opet izvrsava identican upit kao malopre ali sada dobija drugaciji rezultat, odnosno rezultat koji je upisala druga transakcija u medjuvremenu.



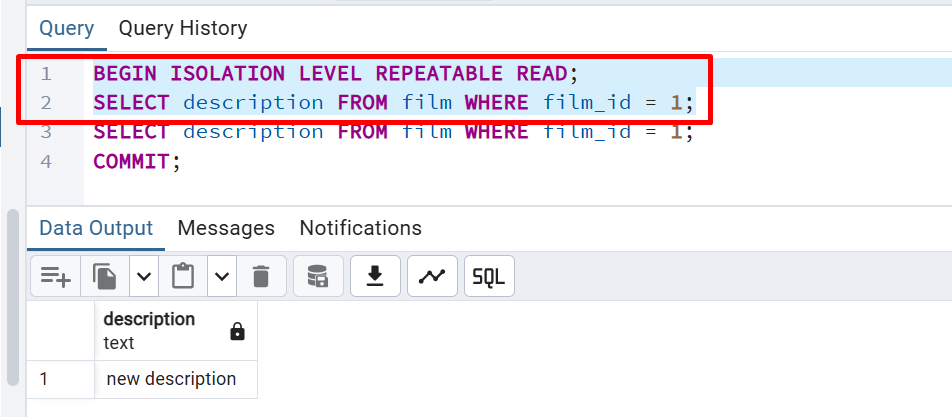
To nije pozeljno ponasanje i predstavlja anomaliju non-repetable read.

## Repeatable read

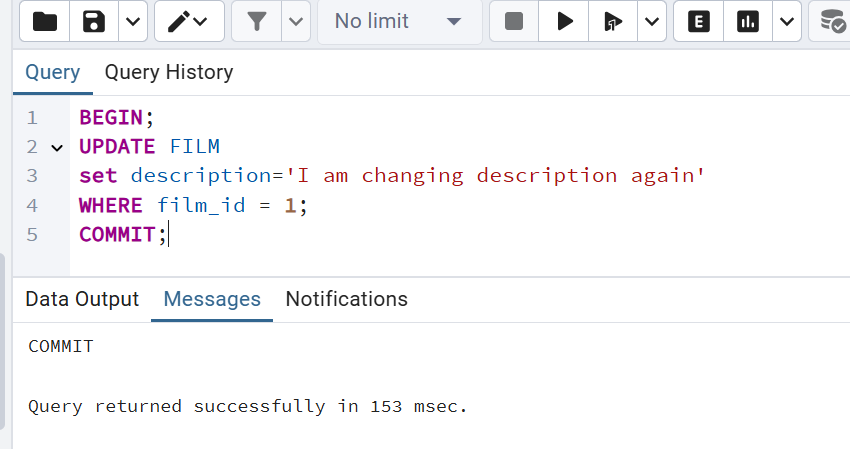
Kao i u prethodnom nivou sprecen je dirty read ali je dodatno **sprecen non-repetable read i lost update**. Nijedna druga transkacija ne moze da menja podatke koji se citaju u trenutnoj transakciji dok se ona ne zavrsi. Garantuje da u okviru jedne transakcije sva ponovljena citanja vracaju iste podatke.

Primer prethodnih transakcija ali na izolacionom nivou repetable read

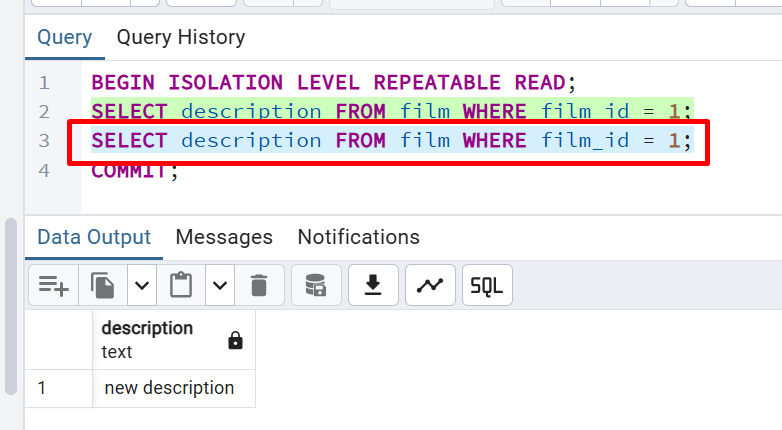
Prva transakcija je pocela sa izvrsavanjem, i izvrsio se samo oznaceni deo, vidimo da je upit vratio trenutni podatak:



U medjuvremenu druga transakcija se izvrsava u celosti i menja bas taj podatak koji prva transakcija cita:



Ali ovog puta kada prva transakcija nastavi sa radom dobija isti rezultat koji je videla na pocetku izvrsavanja:



Primecuje se da je na ovom nivou izolacije sprecen non-repetable read.

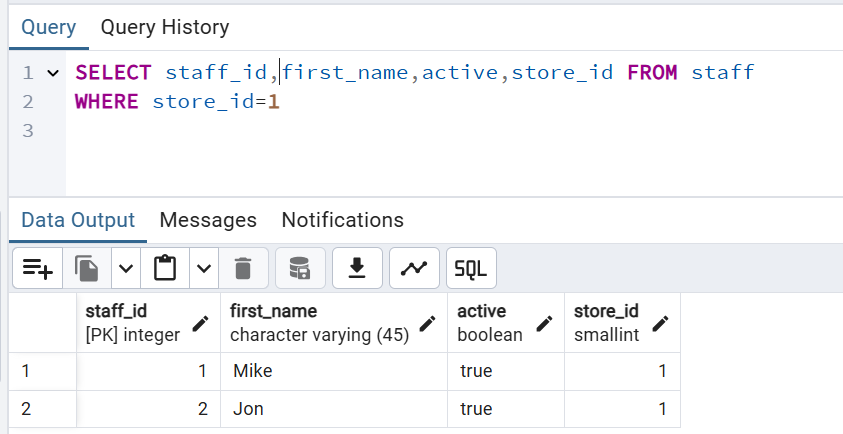
U nekim bazama moze da se desi phantom read na repetable read nivou izolacije, ali u bazama koje implementiraju MVCC, kao sto je i PostgreSQL, phantom read je sprecen zahvaljujuci snapshot mehanizmu. PostgreSQL koristi MVCC mehanizam pa svaka transakcija vidi konzistentan snapshot baze u trenutku kada je ta transakcija pocela (detaljnije u 4.2. poglavlju).

Write skew je jedina anomalija koja je dozvoljena na nivou repetable read, ostale anomalije su sprecene.

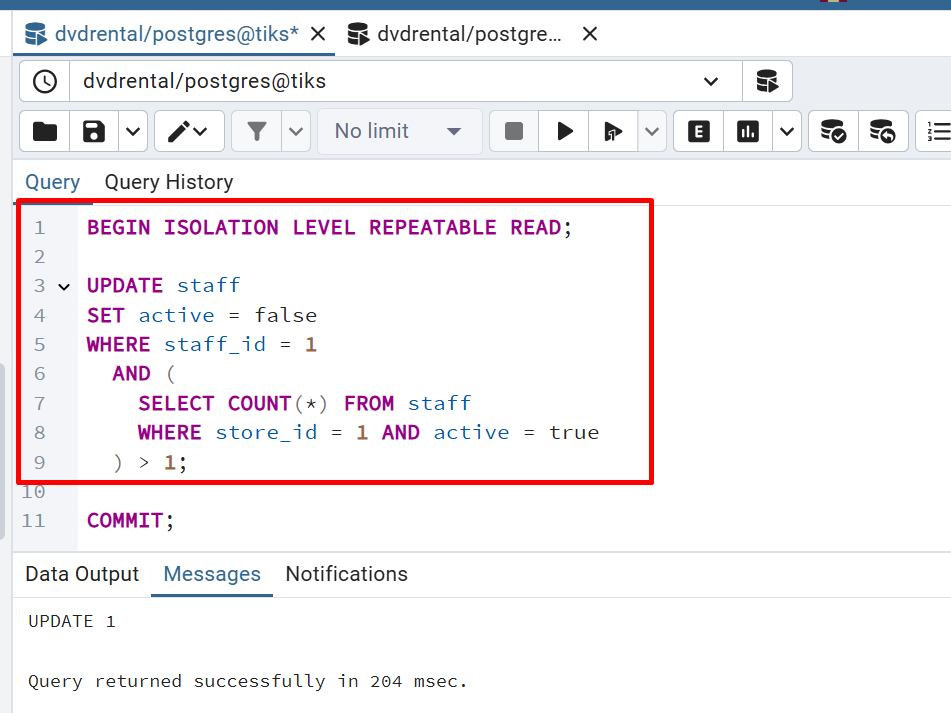
Primer write skew:

Postoje dve transakcije i obe proveravaju da li u radnji postoji vise od 1 aktivnog radnika, ako postoji onda jedan radnik moze zavrsiti smenu, odnosno stavlja se active=false.

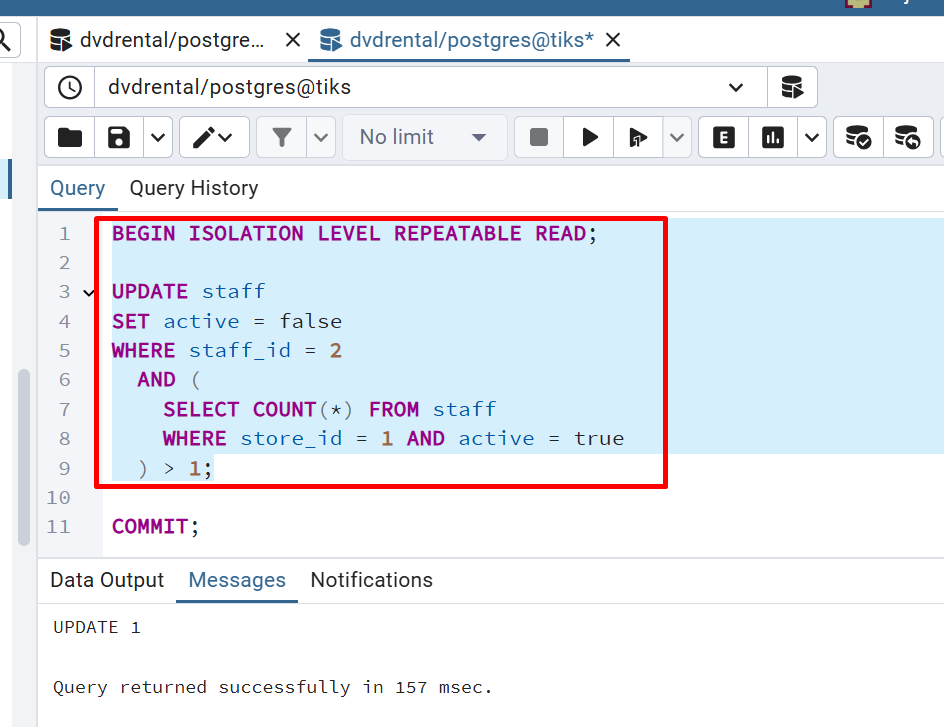
Pre pocetka obe transakcije postoje dva aktivna radnika, a pretpostavimo da je potrebno da samo jedan radnik bude aktivan u jednom trenutku



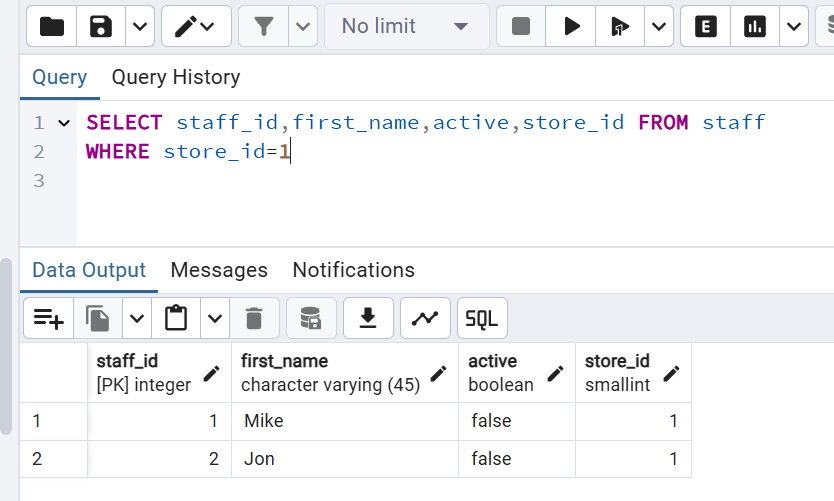
Prva transakcija proverava da li je broj aktivnih radnika veci od 1 u toj radnji, vidi da jeste i izvrsava se. Tacnije radniku sa staff\_id=1 postavlja kolonu active na false i zavrsava mu smenu.



Druga transakcija istovremeno proverava to isto, takodje vidi da ima vise od jednog radnika u radnji ali postavlja active=false radniku sa staff\_id=2.



Nakon commitovanja obe transakcije, ukoliko se opet izvrsi pocetni upit za proveru stanja u bazi:



Iako nije doslo do fizickog konflikta izmedju transakcija, svaka transakcija je menjala drugi red u tabeli, doslo je do logickog konflikta koji ugrozava integritet podataka. Za oba radnika je postavljeno active=false sto ne sme da se desi jer ostavlja bez nadzora radnju koja radi 24/7.

## Serializable

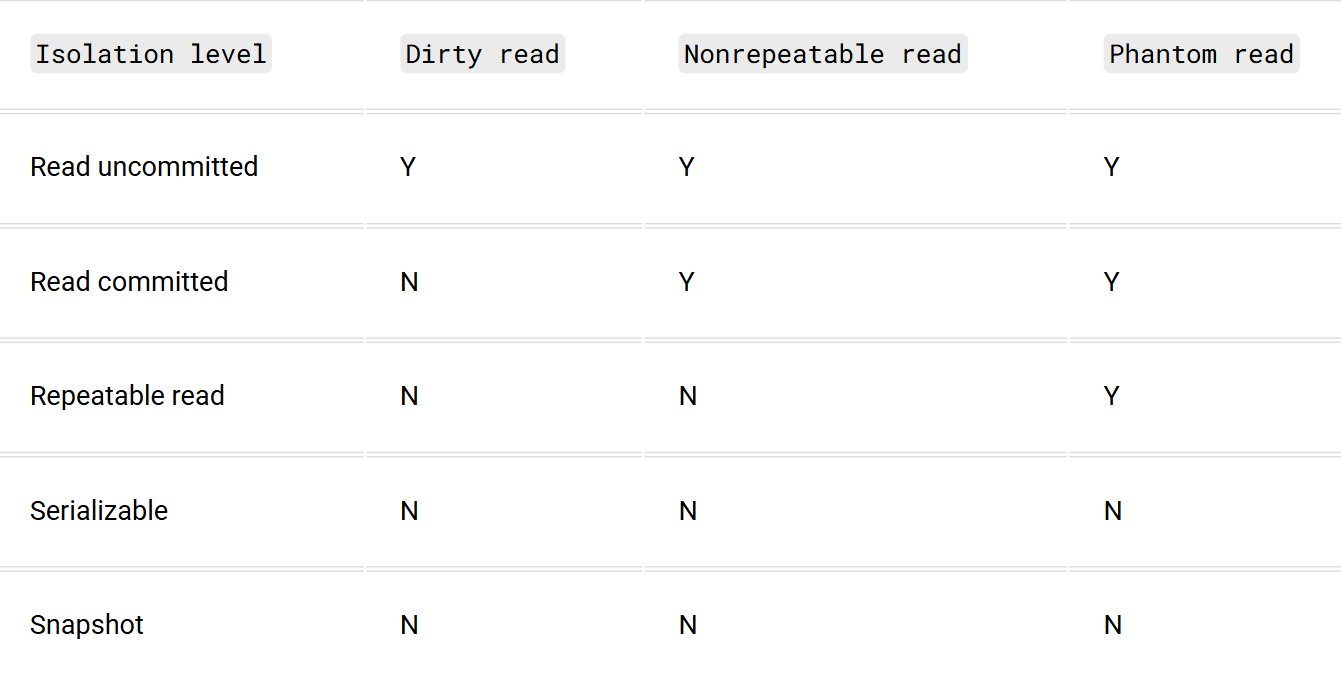
Transakcije se izvrsavaju kao da su serijalizovane jedna za drugom, iako se vremenski preklapaju. Transakcija ne pocinje sa izvrsavanjem dok se trenutna transakcija u potpunosti ne zavrsi. Reseni su sve anomalije izolacije. Negativna strana je sto su ovde performanse najslabije u poredjenju sa ostalim nivoima izolacije, zato sto transakcija ceka da se trenutna transakcija zavrsi cak i kada bi obe trasanakcije mogle da se izvrse paralelno bez konflikata.

# Primeri anomalija zbog kojih je izolacija potrebna

**Race condition** je situacija u konkurentnim sistemima kada vise procesa ili transakcija pristupa nekim podacima a krajnji rezultat zavisi od redosleda izvrsavanja tih procesa ili transakcija. Ako se procesi izvrsavaju u neocekivanom redosledu krajnji rezultat moze biti pogresan ili nekonzistentan. Izolacija transakcija sluzi da spreci ili detektuje race conditione cime se obezbedjuje konzistentnost podataka, odnosno da je krajnji rezultat transkacija predvidljiv i da ne zavrsi od toga koja transakcija ce se prva izvrsiti.

Tipicni pimeri race condition-a su:

* Dirty read
* Dirty write
* Non-repetable read
* Phantom read
* Lost update



Na slici iznad je tabela sa nivoima izolacije i da li je moguce da se jave odredjene anomalije u njima.

*Dirty read scenario*

Pretpostavimo za dve transakcije pristupaju tabeli customer istovremeno.

Prva transakcija je:

BEGIN

UPDATE customer

SET email = 'newemail@gmail.com'

WHERE customer\_id = 1;

Druga transakcija je:

SELECT email

FROM customer

WHERE customer\_id = 1;

Kod dirty read-a bi druga transakcija videla novi email koji upisuje prva transakcija iako ga ona jos uvek nije potvrdila. Ukoliko prva transkacija odradi rollback, to bi znacilo da je transakcija 2 procitala nevazece podatke.

Ovo je primer dirty read-a ali nije moguce da se dirty-read pojavi u PostgreSQL-u zbog njegovog MVCC sistema.

*Non-repetable read scenario*

Ukoliko imamo transakciju 1

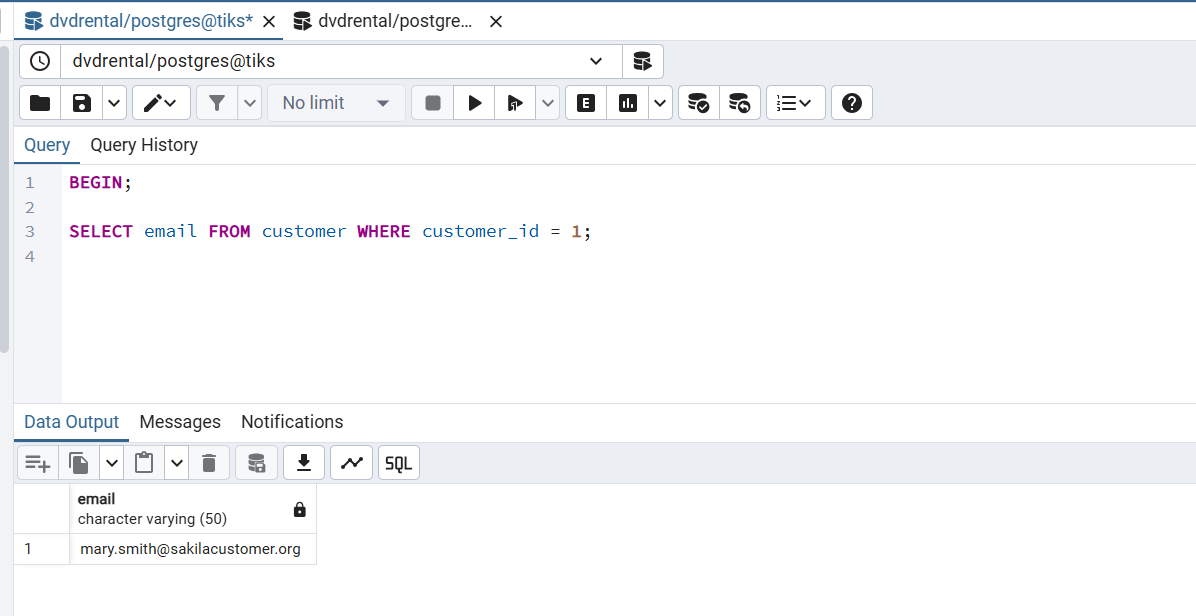
BEGIN;

SELECT email FROM customer WHERE customer\_id = 1;

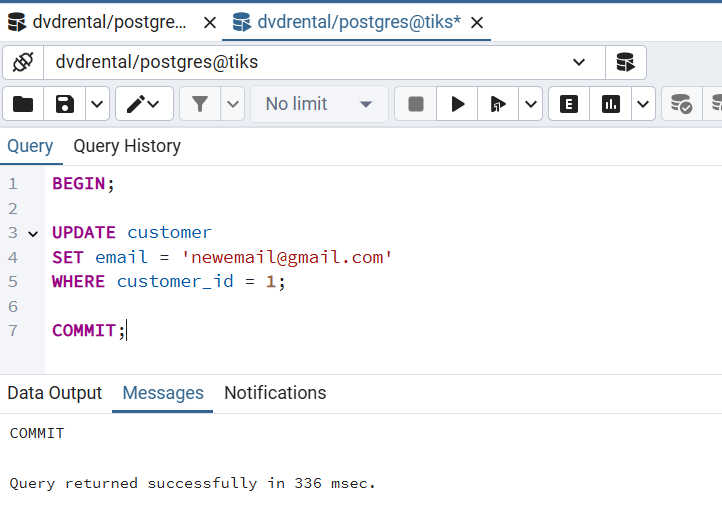
…

SELECT email FROM customer WHERE customer\_id = 1;

Transakcija 1 je pocela sa izvrsavanjem i izvrsio se deo transkcije sa slike ispod, vidi se da je vracen email customera sa customer\_id=1

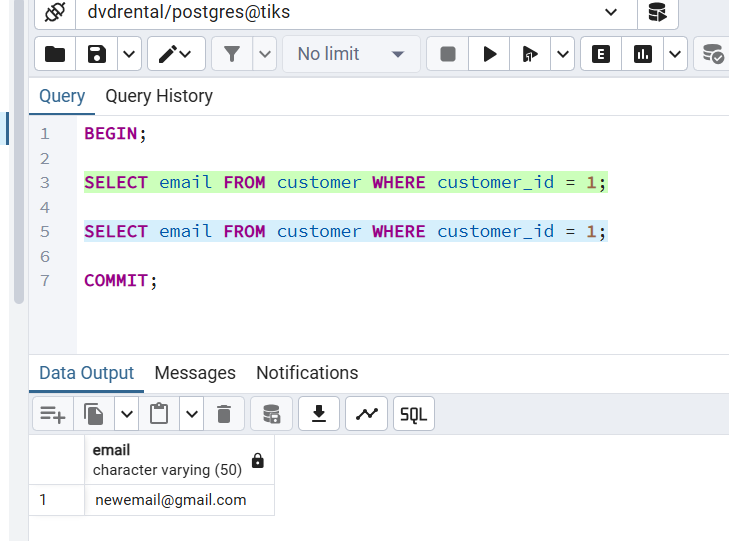


A u medjuvremenu transkacija 2 krece i je izvrsena u potpunosti:



Transkacija 2 je izmenila onaj podatak koji transkacija 1 cita.

Sada kada transakcija 1 nastavlja sa izvrsavanjem i opet izvrsava identican upit ali dobija razlicit rezultat.

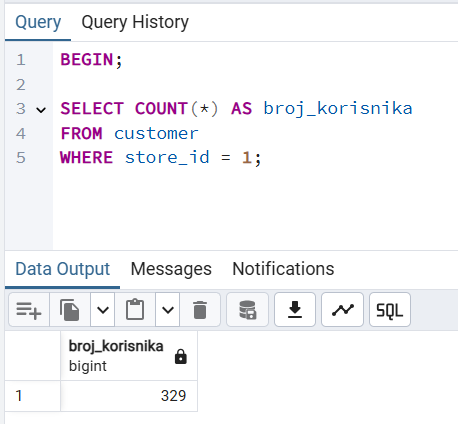


Ova anomalija se zove non-repetable reads, zato sto jedan te isti upit vraca 2 razlicita rezultata u okviru iste transkacije.

*Phantom-read scenario*

Ukoliko transkacija 1 procita skup redova koji zadovoljaju neki uslov,druga transkacija u medjuvremenu ubaci nove redove koji zadovoljaju taj isti uslov, a transkacija 2 opet cita te redove, desice se da transakcija 1 vidi dodatne redove.

Transakcija 1 je pocela sa izvrsavanjem i izvrsio se ovaj deo transkacije:



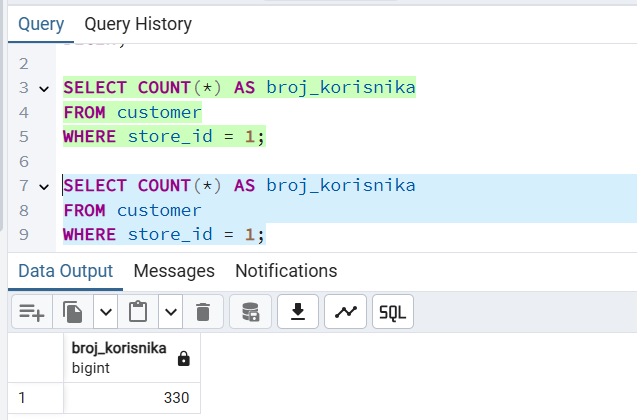
Vidi se da je broj korisnika koji vidi transkacija 1 trenutno 329.

U medjuvremenu se izvrsila kompletna transakcija 2:



Transakcija 1 nastavlja za izvrsenjem i opet izvrsava identican upit ali sada dobija drugaciji rezultat

Odnosno, sada postoji dodatan red koji zadovoljava uslov upita. To je phantom-read.



# Mehanizmi za obezbedjivanje konkurentnog pristupa

U PostgeSQL-u mehanizmi koji implementiraju razlicite nivoe izolacije se zasnivaju na kombinaciji MVCC i dodatnih mehanizama zakljucavanja.

## MVCC

Multi-Version Concurrency Control je osnovni mehanizam koji PostgreSQL automatski koristi da bi omogucio konkurentni pristup bazi podataka bez zakljucavanja za citanje. Svaka transkacija vidi snapshot baze koji je vazio u trenutku kada je ta transakcija pocela. Koristi se da bi baza podrzala visoke nivoe izolacije, poput Repetable read i Serializable.

Osnovna ideja MVCC je da umesto da transakcije blokiraju jedna drugu prilikom citanja i pisanja, da se svaka transakcija izvrsava i da svaka transakcija stvara novu verziju reda u tabeli. Svaki red u tabeli sada ima i dodatne kolone **xmin** i **xmax**.

Xmin predstavlja ID transakcije koja je kreirala red, a xmax predstavlja ID transkacije koja je obrisala ili azurirala red. Kada transakcija cita red proverava se vrednost xmin i xmax. Ako je xmin manji od ID trenutne transakcije a xmax je veci od ID transakcije ili je NULL onda je red vidljiv za tu transakciju. U suprotnom se taj red ignorise (nije deo snapshota). Na ovaj nacin svaka transkacija vidi konzistentan snapshot baze cak i kada druge transakcije vrse promene. Transakcije mogu paralelno da citaju podatke bez cekanja drugih transakcija.

## Snapshot mehanizam

Snapshot je mehanizam rada koji koriste moderne baze zasnovane na MVCC (Multi Version Concurrency Control) principu, ukljucujuci i PostgreSQL.

Upit moze da koristi podatke samo ako ce baza biti u konzistentnom stanju u toku cele transakcije. Ako neka druga transakcija izmeni podatke dok se oni jos koriste u trenutnoj transkaciji, to trenutna transkacija ne vidi. Trenutna transakcija ce raditi sa snapshotom podataka koji su postojali na pocetku transakcije. Snaphost transakcije ne zahteva zakljucavanje podataka prilikom citanja podataka niti zabranu ostalim transkacijama da menjaju te podatke. S obzirom da transakcije ne blokiraju citanja i upise, vreme cekanja se smanjuje u odnosu na serializable nivo. Medjutim, moze doci do konflikata pri pisanju (write-write conflict) koji se resavaju abortovanjem jedne transkacije. Takodje moze doci i do write screw-a. Write screw nastaje kada razlicite transakcije menjanju razlicite redove ili kolone u bazi ali zajednicki rezultat tih transakcija krsi integritet podataka.

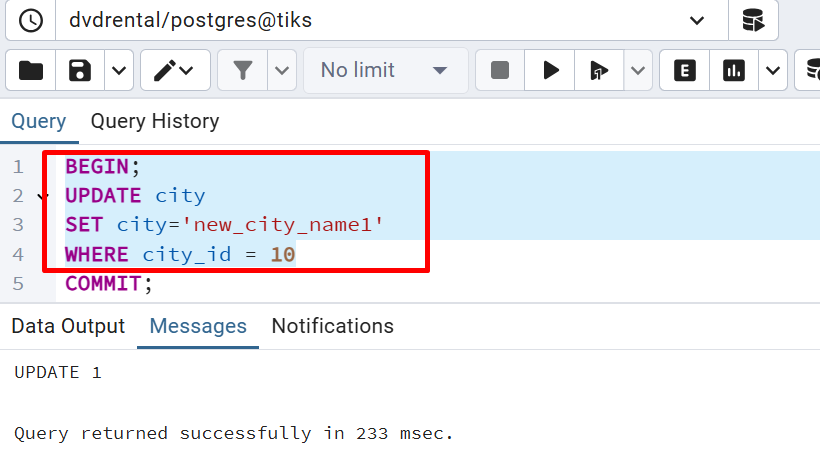
## Row locking

Row locking je osnovni mehanizam kontrole konkurentnosti u PostgreSQL-u. Omogucava da vise transakcija istovremeno pristupa istoj tabeli, ali razlicitim redovima. Ukoliko razlicite transakcije pokusaju da izvrse operacije UPDATE ili DELETE nad istim redovima iste tabele dolazi do blokiranja transakcija. Omogucava paralelne upise u razlicite redove iste tabele i predstavlja balans izmedju izolacije i performansi baze podataka.

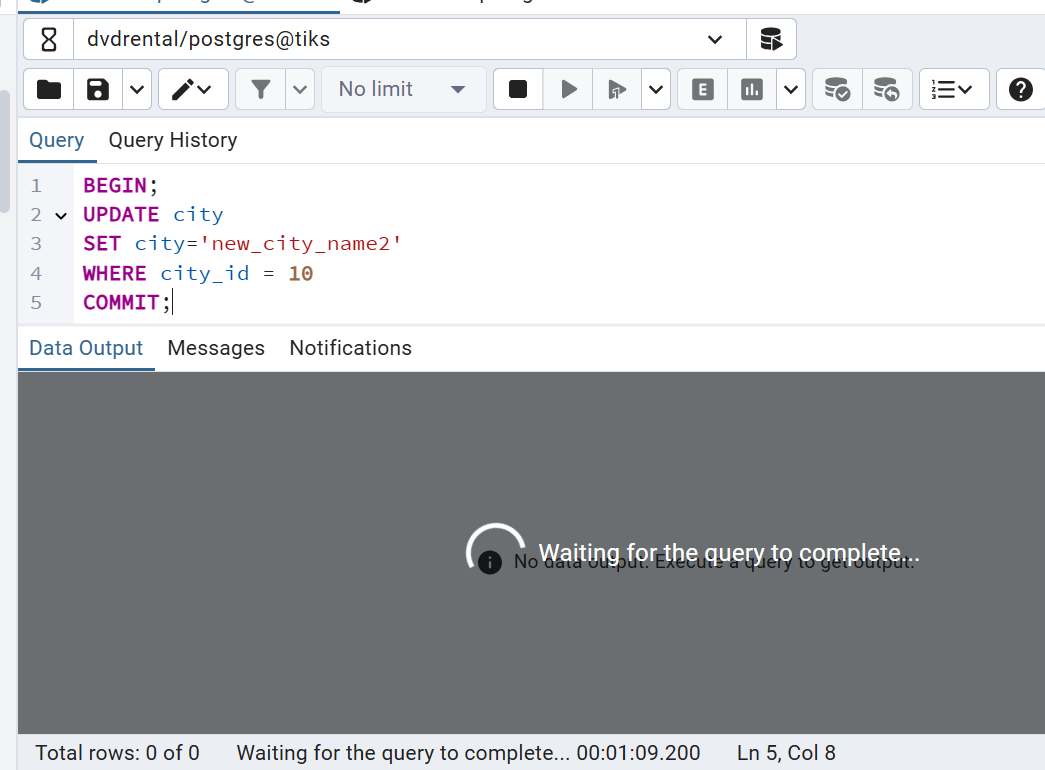
Kada transakcija izvrsava komandu za upis podataka u tabelu, PostgreSQL automatski zakljucava te redove tako da druge transakcije ne mogu da ih promene, ali mogu da ih citjau.

Primer sa 2 paralelna upisa:

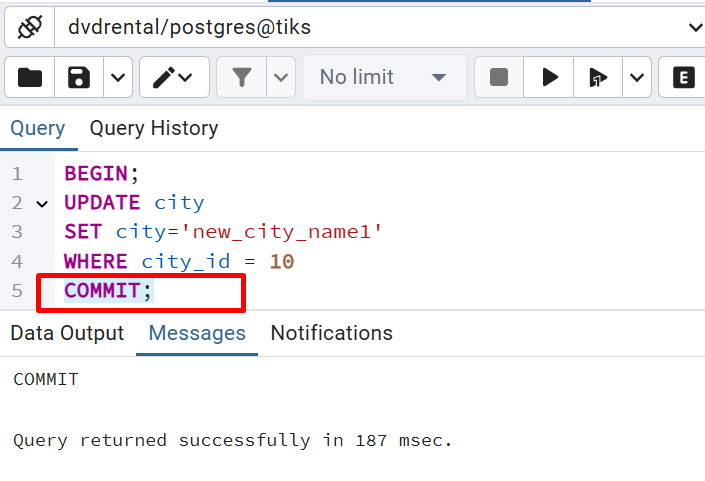
Trenutna transakcija je izvrsila oznaceni deo:

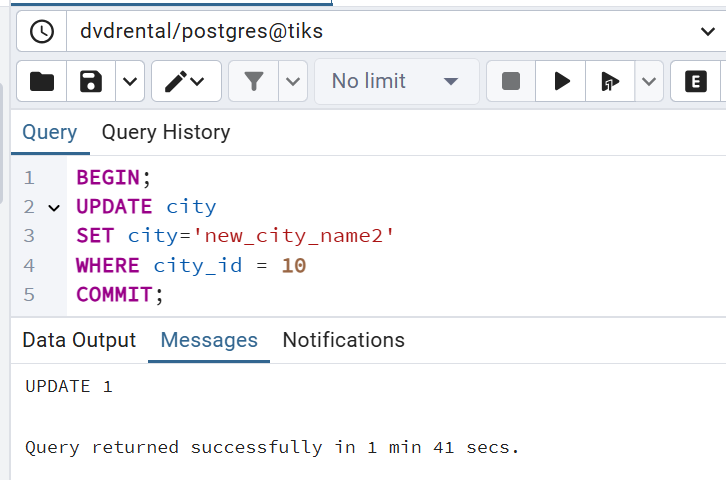


U isto vreme druga transakcija pokusava da se izvrsi u celosti i da izmeni isti red, red sa city\_id=10, ali ne moze da pristupi tom redu vec ceka dok se ne zavrsi prva transakcija.



Tek kada se izvrsi commit(ili rollback) iz prve transakcije, ova transakcija moze da nastavi sa izvrsenjem, jer COMMIT oznacava da se prva transakcija zavrsila.



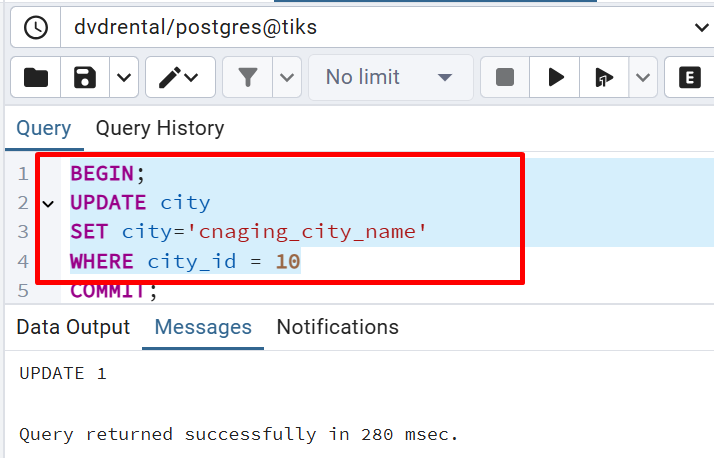


Vidimo da je vreme izvrsavanja skoro 2 minuta, uzrok toga je sto je transakcija cekala da se zavrsi transakcija koja je prva okupirala trazeni red.

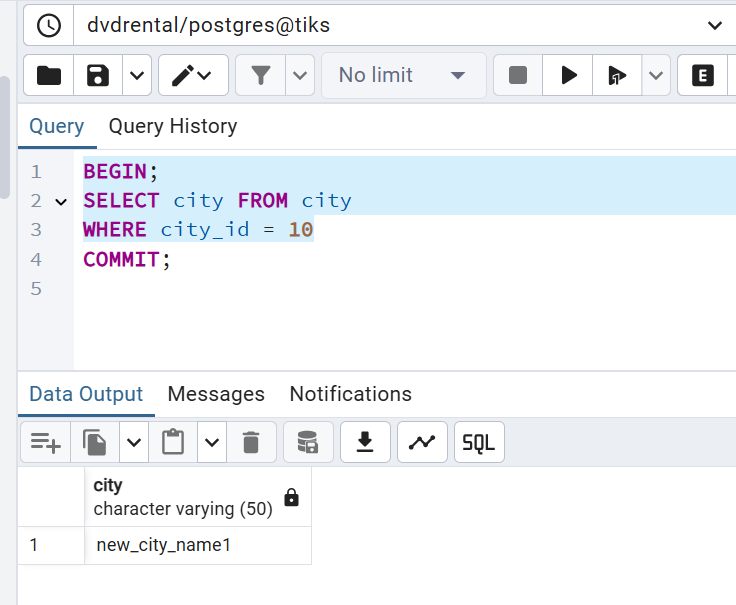
Primer sa jednim upisom i jednim citanjem:

Kada transakcija upisuje u red, ona taj red zakljucava samo za upis drugim transakcijama. Druge transakcije mogu da citaju podatke iz tog reda ali nece dobiti podatak koji transakcija trenutno menja, jer nije commitovan, vec ce dobiti poslednji commitovani podatak.

Prva transakcija izvrsava oznaceni deo:



Druga transakcija cita taj podatak, istovremeno dok traje prva transakcija:



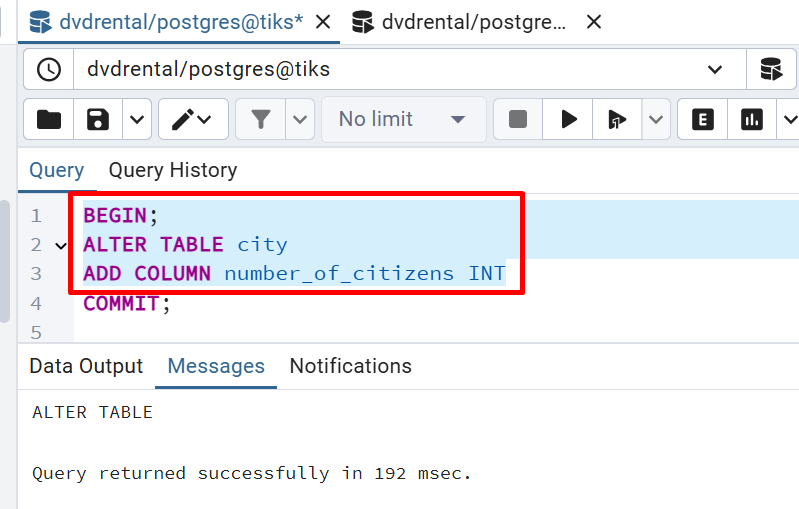
Primecuje se da je transakcija procitala poslednji commitovani podatak iz ovog reda.

## Table locking

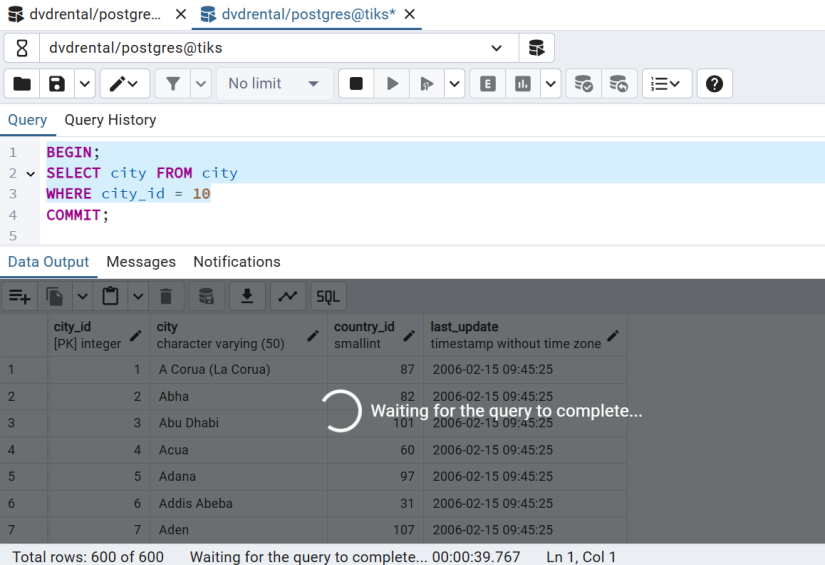
Zakljucavanje obuhvata citavu tabelu, ukljucujuci njene redove, indekse i metapodatke. Zakljucavanje na nivou tabele se koristi redje nego zakljucavanje na nivou redova. Koristi se kada se izvrsavaju operacije nad tabelom kao sto su ALTER TABLE ili DROP TABLE. PostgreSQL automatski primenjuje table locking prilikom izvrsavanja tih operacija. Kada se cela tabela zakljuca ostale transakcije koje pokusavaju da menjaju podatke iz te tabele ce biti blokirane sve dok se transakcija ne zavrsi. Ovo obezbedjuje konzistentnost podataka ali i moze narusiti performanse zato sto smanjuje nivo paralelizma.

Primer:

Prva transakcija izvrsava:



Druga transakcija ne moze da izvrsi cak ni citanje podataka iz te tabele dok se ne zavrsi transakcija koja menja tabelu.



Razlog za to je sto PostgreSQL kod menjanja tabela automatski koristi Acess Exclusive Lock, sto je najstrozi tip zakljucavanja tabele, i blokira sve druge pristupe.

## Predicate locking

Predicate locking je oblik zakljucavanja koji ne zakljucava konretne redove u tabeli vec zakljucava uslov koji se definise nad tabelom. Transakcija zakljucava sve redove koji zadovoljvaju neki uslov i sve buduce redove koji zadovoljavaju taj uslov. Ukoliko neka transakcija cita podatke koji zadovoljavaju odredjeni uslov i stavi lock nad tim uslovom, a neka druga transakcija pokusa da doda red koji bi odgovarao tom uslovu, prva transakcija ce i dalje videti nepromenjene podatke sve dok se ne zavrsi. Na ovaj nacin se sprecavaju phantom reads.

# Zakljucavanje i upravljanje konkurentnim pristupom

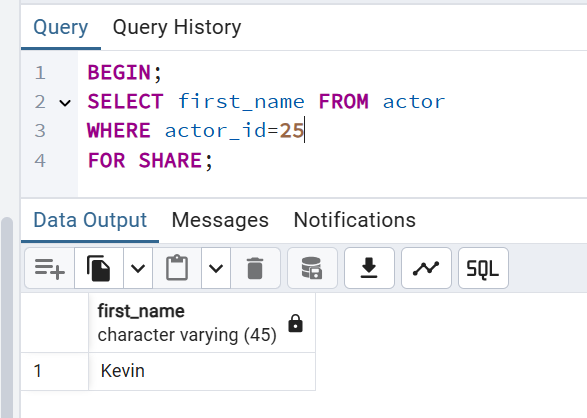
Jedan od najvaznijih mehanizama kojima se odrzava kontrola konkurentnosti u bazama podataka gde vise korisnika istovremeno pristupa podacima su **lock-ovi(zakljucavanje)**. Zakljucavanje predstavlja nacin da baza kontrolise pristup odredjenom resursu kako bi se srpecilo da istovremeno izvrsavanje transkacija dovede do nekonzistenih rezultata ili gubitka podataka.

## Pessimistic locking

Pessimistic locking se u PostgreSQL-u se mora upotrebiti rucno(implicitno). Transakcija zakljucava red koji cita ili menja, kako bi druge transakcije cekale dok se trenutna transakcija ne zavrsi.

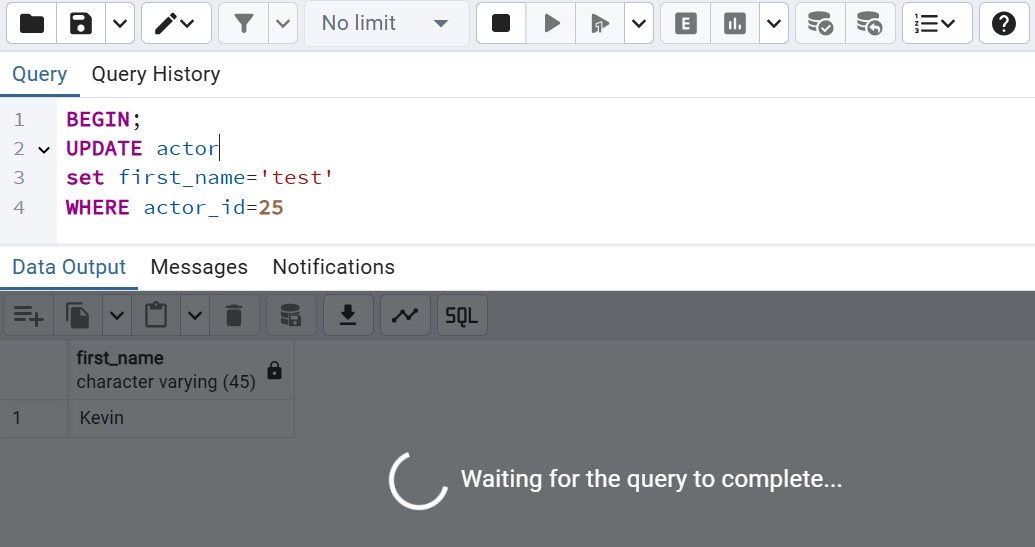
Pessimistic locking pretpostavlja da ce do konflikata izmedju transakcija dolaziti cesto. Kada transakcija pristupi podacima za citanje ili pisanje, baza odmah zakljucava te podatke. Ostale transakcije koje zele da pristupe tim podacima moraju da cekaju dok se trenutna transkacija ne zavrsi da bi se podaci otkljucali. Koristi se u sistemima gde je konzistentnost vrlo vazna i prihvatljivo je da se transakcije duze izvrsavaju kao sto su bankarski sistemi, sistemi za rezervacije. Lockovi koje pessimistic locking koristi mogu da budu Exclusive lock i Shared lock.

**Shared Lock** se koristi kada transkacija samo cita podatke iz baze ali ih ne menja. Vise transkacija moze istovremeno drzati shared lock nad istim redom ili tabelom s obzriom da ctanje ne remeti konzistentnost podataka. Sintaksom **SELECT … FOR SHARE** se stavlja Shared lock na procitane redove.

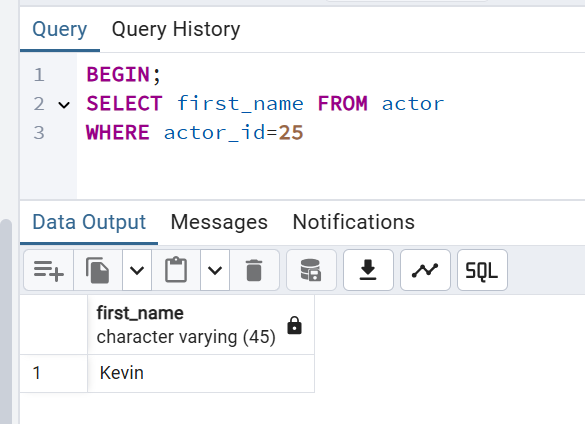


Ovom transakcijom je red iz tabele actor sa actor\_id=25 shared zakljucan i ostale transakcije mogu da citaju taj podatak ali ne mogu da ga menjaju sve dok se ne izvrsi COMMIT iz trenutne transkacije.

Transakcija koja zeli da azurira taj podatak mora ceka u redu i izvrsava se tek kada se trenutna transkacija zavrsi.



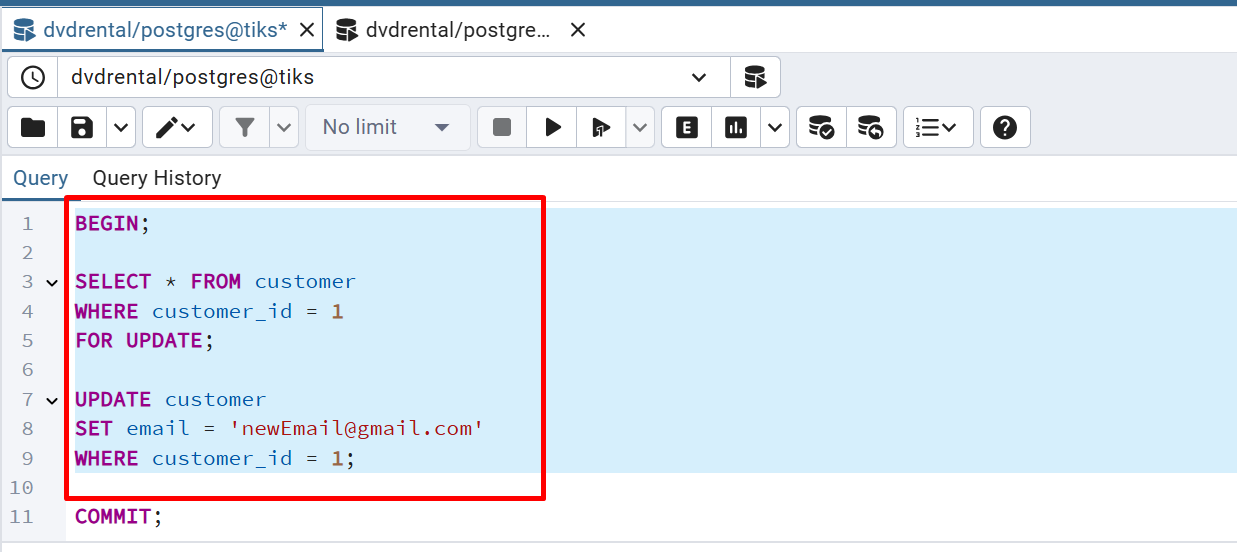
Transkacija koja zeli samo da procita taj podatak to moze da uradi bez cekanja.



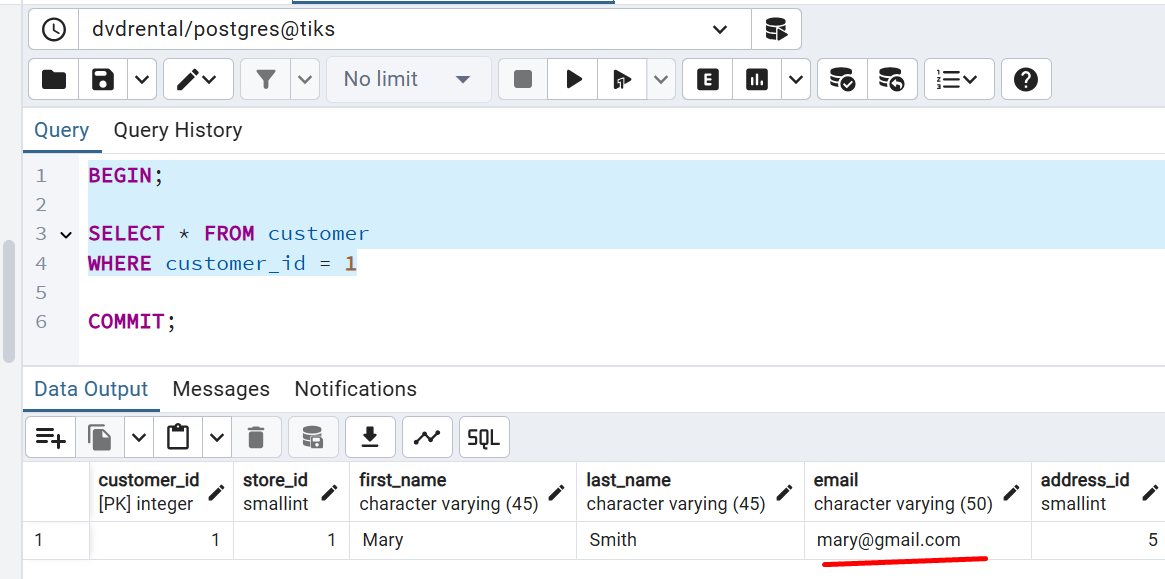
**Exclusive lock** se koristi kada transakcija menja podatke iz baze. Da ne bi dolazilo do nekonzistentnosti neophodno je da samo jedna transakcija ima pristup podacima koji se menjaju. Dok traje X-lock nijedna druga transkacija ne moze dobiti ni X-lock ni S-lock nad tim podacima. Ipak, obican SELECT se moze izvrsiti, s tim sto ce se citati poslednji commitovani podaci, pre pocetka transkacije koja trenutno drzi X-lock.

Primer

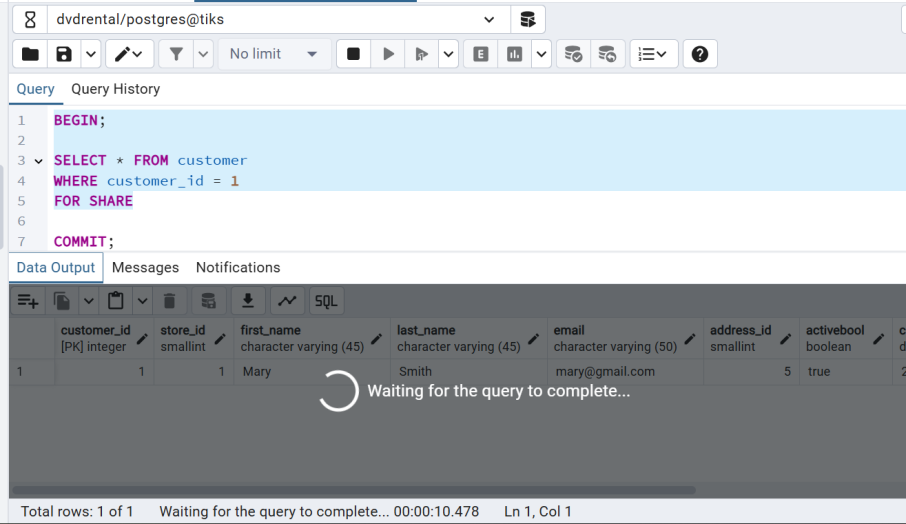
Kada prva transakcija izvrsi:



Druga transakcija cita podatke koji poslednji su commitovani:



Ipak, ako prva transakcija i dalje drzi X-lock a druga transakcija hoce da stavi S-lock ili X-lock na te podatke, ona mora da ceka da se prva transakcija zavrsi:



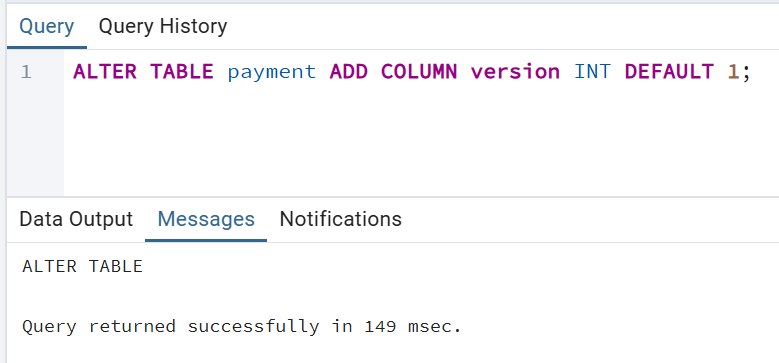
## Optimistic locking

Optimistic locking ne koristi fizicke lockove. Optimistic locking se kao i Pessimistic lokcing mora upotrebiti rucno. Transakcija ne zakljucava red dok ga cita ili menja, vec proverava verziju redova prilikom komitovanja. Ako se verzije razlikuju u odnosu na verzije sa pocetka transakcije, transakcija se abortira.

Optimisitic lockingpretpostavlja da konflikti izmedju transakcija nastaju retko. Transakcija ne zakljucava podatke unapred kada ih cita. Kod ovog pristupa je prioritet konkurentnost a konflikti se resavaju pri upisu.

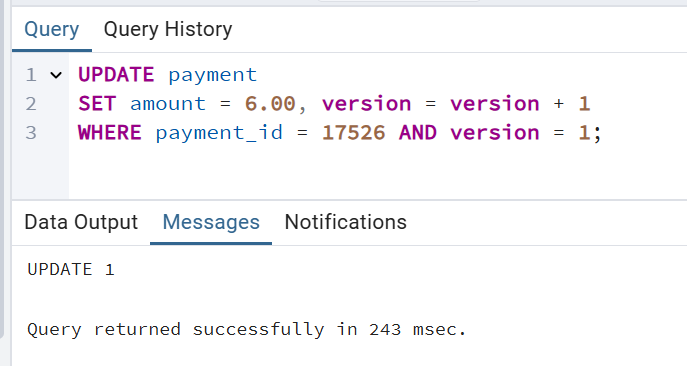
Osnovna ideja optimistic lockinga je da se svakoj tabeli ili zapisu u bazi doda verziona kolona. Najcesce je to celobrojni brojac (version) ili vremenska oznaka (timestamp).

Kada aplikacija ucita neki red iz baze podataka, ona pamti i vrednost verzione kolone. Kada je potrebno da se taj red azurira aplikacija proverava da li je trenutna vrednost verzione kolone ista kao i kada se red ucitao. Ako je vrednost ista to znaci da taj red nije menjan u medjuvremenu i moze se jednostavno izvrsiti azuriranje. Ako je vrednost razlicita to znaci da je u medjuvremenu neka druga transakcija izmenila taj red i trenutna transakcija mora da se ponisti i da se izvrsi ispocetka sa novom verzijom tog podatka. Dodavanje kolone version u tabeli payment koja ce sluziti za kontrolu verzija:



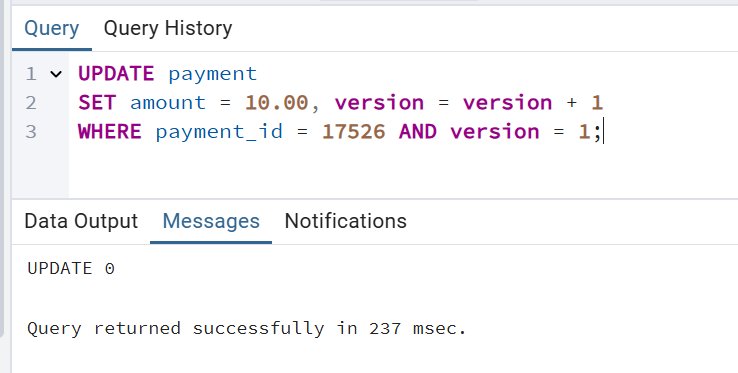
Pretpostavimo da postoje dve paralelne transkacije koje obe menjaju isti red iz tabele payment.

Prva transakcija se uspesno izvrsila i azurirala je kolonu version, odnosno povecala je za 1.



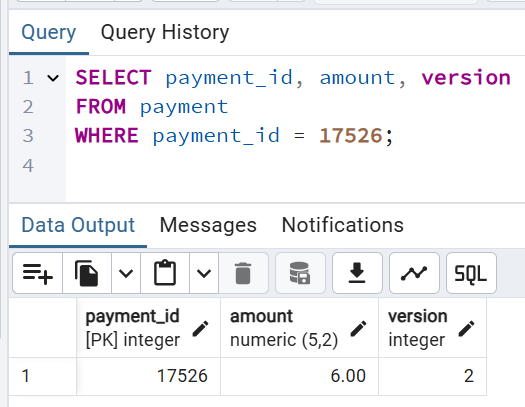
UPDATE 1 znaci da je jedan red azuriran.

Druga transakcija pokusava da azurira taj isti red ali i dalje koristi version = 1 sto je zastarela verzija. Baza ne pronalazi odgovarajuci red, s obzirom da kolona version nije vise 1 nego 2.

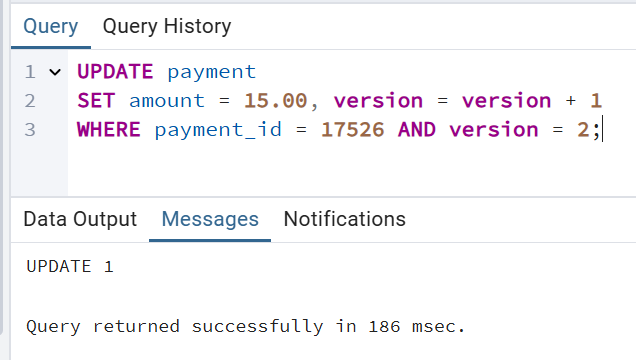


UPDATE 0 znaci da nijedan red nije azuriran.

Transakcija 2 mora da prvo ucita red sa payment\_id = 17526 da bi dobila podatak o trenutnoj verziji a tek onda da azurira taj red.



Ukoliko se sada izvrsi query sa version = 2, baza pronalazi odgovarajuci red i azurira ga:



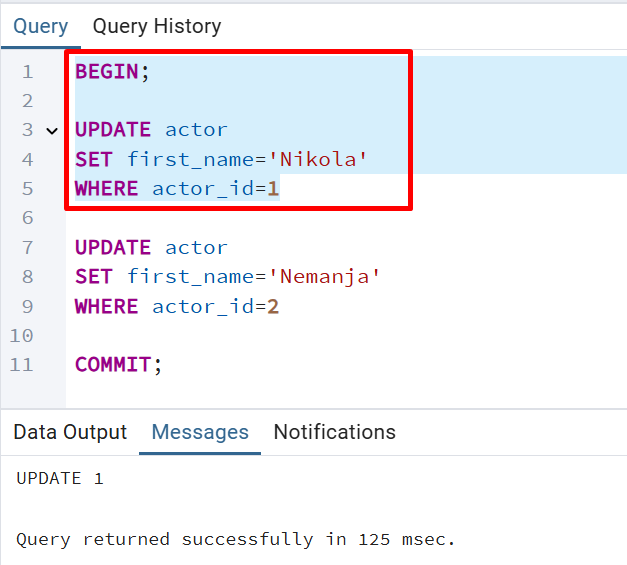
Sada je UPDATE 1, sto znaci da je jedan red je azuriran.

## Problemi u upravljaju zakljucavanjem

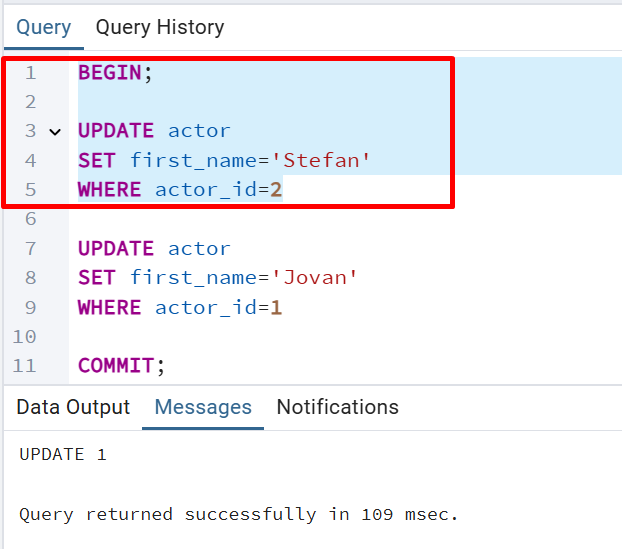
**Deadlock** se desava kada su dve ili vise transakcija medjusobno blokirane zato sto cekaju jedna drugu i nijedna transakcija ne moze da nastavi sa izvrsavanjem dok je druga transakcija ne odblokira a to se nece desiti nikada. Deadlock obicno nastaje kada transakcije zakljucavaju iste resurse u razlicitom redosledu.

Primer:

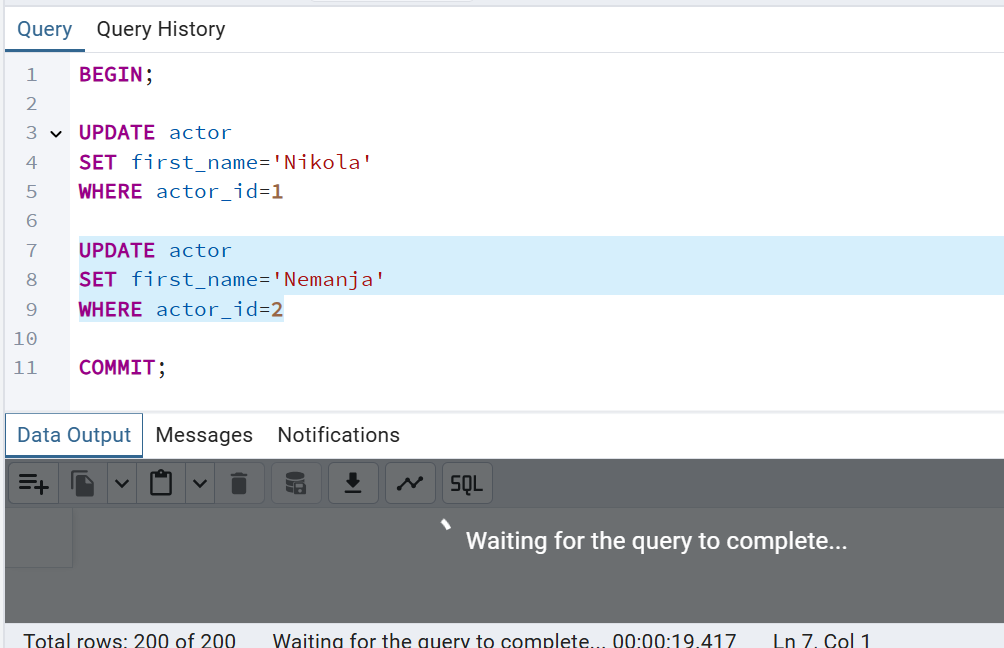
Prva transakcija je za sada izvrsila deo:



Transkaciju 2 je za sada izvrsila deo:



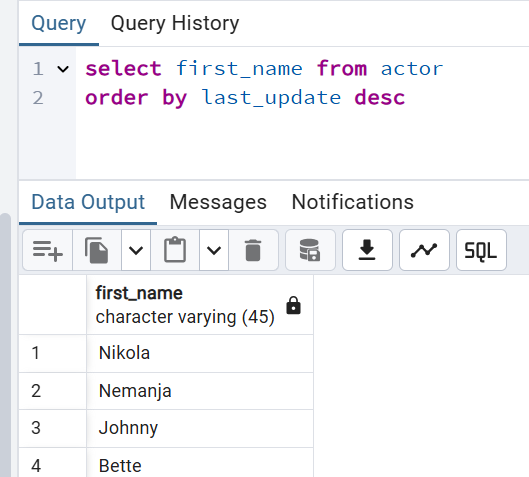
U ovom slucaju, kada transakcija 1 pokusa da pristupi sledecem upitu, odnosno kada pokusa da azurira red kojem je actor\_id=2 ona mora da ceka da ga transakcija 2 oslobodi zato sto ona trenutno drzi lock na njemu.



Transakcija 2 da bi ga oslobodila mora da se izvrsi u potpunosti, odnosno da se izvrsi naredba COMMIT, medjutim transkacija 2 se nece izvrsiti do kraja zato sto ona u isto vreme ceka da transakcija1 oslobodi red kojem je actor\_id=1



PostgreSQL ima **deadlock detektor** i ova poruka govori da je PostgreSQL primetio da trenutno postoji deadlock. Transakcija koja je mladja, odnosno ona koja je kasnije krenula sa izvrsavanjem se prekida(abort) da bi starija mogla da nastavi sa izvrsavanjem. Krajnji rezultat ove dve transkacije je ustvari rezultat samo transkacije 1, transakcija 2 je prekinuta.

Prevencija deadlocka:

* Dosledan redosled zakljucavanja - sve transakcija zakljucavaju resurse u istom redosledu
* Kratke transakcije - sto krace traje transkacija to su manje sanse za deadlock
* Minimalno zakljucavanje - zakljucati samo one podatke koji su potrebni
* Optimistic locking moze smanjiti pojavu deadlocka
* Koristiti Shared lock ukoliko je moguce, a Exslusive lock koristiti samo ako je neophodan
* Postaviti timeout za cekanje na lock

**Lock escalation** je proces u kojem baza podataka automatski konvertuje vise manjih zakljucavanja (uglavnom redova) u jedno veliko zakljucavanje (uglavnom tabelu). Desava se kada je veliki broj redova iste tabele zakljucan a to iziskuje veliku potrosnju resursa jer svaki poseban lock zauzima deo memorije i procesora. Baza tada radi lock escalation odnosno umesto da drzi mnogo manjih redova onda ih spaja u jedan veliki lock i jednim lockom drzi celu tabelu.

**Livelock** je situacija kada transakcije ne napreduju, ali za razliku od deadlocka one ne cekaju, vec non stop nesto rade ali bez pravog pomaka sa izvrsenjem. Transakcije zauzimaju procesor i memoriju ali bez ikakvog pozitivnog efekta na napredak sistema.

**Starvation** je situacija u konkurentnim sistemima kada neki proces ili transakcija nikada ne dobija pristup resursu koji mu je potreban iako je taj resurs dostupan i druge transakcije normalno rade. Javlja se zbog loseg rasporeda izvrsavanja procesa, kada odredjeni procesi stalno imaju prioritet u odnosu na ostale.

## Zakljucavanje i performanse

Konkurentnost u bazama podataka oznacava mogucnost da vise transakcija istovremeno pristupa bazi, sto je vece ogranicenje odnosno veci stepen izolacije to je utrosak resursa veci.

Visok nivo izolacije(Serializable) moze da spreci sve oblike race condition-a i da garantuje maksimalnu konzistentnost podataka. Ipak, veca izolacija znaci vise cekanja i zakljucavanja transkacija sto moze povecati vreme izvrsavanja.

Nizi nivo izolacija(Read commited) omogucava brze izvrsavanje ali povecavaju rizik od anomalija.

U zavisnosti od tipa aplikacije odredjuje se da li je prioritet na konzistentnosti kao kod finansijskih transakcija (implementiraju se visi nivoi izolacije) ili je prioritet na performansama kao kod sistema sa mnogo citanja i retkim konfliktima (implementiraju se nizi nivoi izolacije). Pozeljno je izabrati najnizi nivo izolacije koji je prihvatljiv za aplikaciju. Dobra praksa je minimizirati kriticnu sekciju sto dovodi do manje zakljucanih redova.

Kljucni faktori konkurentnosti koji uticu na performanse su:

* Optimistic i pessimistic locking - Koristiti optimisticko zakljucavanje ako je moguce, a pesimisticko samo kada je neophodno, odnosno kada su konflikti cesti i kriticni.
* Locks - Odnosi se na vrste pesimistickog zakljucavnja shared i exslusive lock. Koristiti exclusive lock samo kada je neophodan, a kad god je moguce koristi shared lock, jer vise transakcija moze istovremeno da cita podatak sto povecava konkurentnost i smanjuje vreme izvrsavanja.

# Zakljucak

Izolacija i konkurentni pristup podacima su osnova za istovremeni pristup podacima od strane vise korisnika, uz ocuvanje konzistentnosti i integriteta podataka. Kao i u vecini inzenjerskih oblasti, i ovde je neophodno ostvariti odgovarajuci trade-off izmedju performansi sistema i tacnosti rezultata.

Veci nivoi izolacije obezbedjuju vecu konzistentnost podataka ali manju brzinu izvrsavanja transakcija zbog veceg broja zakljucavanja i smanjenja paralelizacije. Nizi nivoi izolacije omogucavaju vecu brzinu izvrsavanja upita ali uz veci rizik od pojave anomalije i nedoslednosti u podacima. Izbor odgovarajuceg nivoa izolacije je kompromis izmedju pouzdanosti i performansi sistema. Mehanizmi zakljucavanja takodje imaju vaznu ulogu za stabilan i predvidljiv rad baze podataka. U sistemima u kojima se ne ocekuje veliki broj istrovremenih pristupa podacima je bolje izabrati optimistic locking zbog boljih performansi, a u sistemima gde se ocekuje cesto pristupanje vise korisnika istim podacima je bolje izabrati pessimisitc locking. Efikasno upravljanje zakljucavanjima sprecava pojavu problema kao sto su deadlock, livelock ili lock escalation koji mogu nastetiti performansama sistema. Pravilno kombinovanje mehanizama zakljucavanja i nivoa izolacija omogucuje da se u zavinosti od prioriteta aplikacije izabere da li je akcenat na brzini izvrsavanja ili konzistentnosti podaka.

# Literatura

1. ,,Transaction isolation levels”, Miriam Antona
2. ,,Topics in Database Management Systems”, Xiangyao Yu
3. ,,Isolation and MVCC”, Egor Rogov, Pavel Luzanov
4. ,,High-Performance Concurrency Control Mechanisms for Main-Memory Databases”, Per-Ake Larson
5. Mastering PostgreSQL 13, Fourth Edition, Hans-Jürgen Schönig
6. <https://www.scribd.com/document/860171708/Isolation-database-systems>
7. <https://www.postgresql.org/docs/current/transaction-iso.html>
8. <https://www.postgresql.org/docs/current/explicit-locking.html>
9. <https://learn.microsoft.com/en-us/sql/relational-databases/sql-server-transaction-locking-and-row-versioning-guide?view=sql-server-ver17>
10. <https://www.scribd.com/document/835805809/Multi-Version-Concurrency-Control-MVCC-in-PostgreSQL>