

# Skalabilnost

## Dizajn šeme baze podataka

Priložen u datoteci model.pdf

## Predlog strategije za particionisanje podataka

Računaćemo da su od 200 miliona korisnika najveći deo pacijenti. S obzirom da aplikacija za klinički centar nema faktor društvenih mreža, odnosno neće postojati korisnici koji su popularniji i zahtevaju veliki broj upita, potencijalno bismo mogli da particionisemo tabelu pacijenata po njihovom id-ju, čime bismo postigli uniformnu raspodelu podataka po serverima. Takođe podaci kojima će se najčešće pristupati su email i lozinka, što nas je dovelo do vertikalnog particionisanja posebne tabele za ta dva obeležja kako bi se pristupalo manjem broju podataka.

S obzirom na veliki broj pristiglih zahteva za pregledima, koje bi administratori klinike odobravali i dodeljivali im sobe, prirodno se nameće particionisanje po id-ju klinike kako bi upiti za određenu kliniku vraćali podatke koji su bliski na serverima. Takođe bi se algoritmi za dodeljivanje soba zahtevima za pregled u slučaju da ih administratori ne odobre mogli podeliti na aplikativne servere gde je svaki zadužen za određene klinike.

Takođe će biti velik broj koji pregleda koji su zakazani ili završeni, koje bismo potencijalno particionisali po vremenu početka, gde i stari pregledi ne bi bili zanemareni jer bi se njima pristupalo radi provere istorije pregleda i istorije bolesti preko zdravstvenog kartona, a budućim pregledima radi provere zakazanih i eventualnog otkazivanja istih.

## Predlog strategije za replikaciju baze i obezbeđivanje otpornosti na greške

S obzirom da će pacijenti mnogo češće pretraživati klinike i/ili doktore ili pristupati svom zdravstvenom kartonu nego što će upisivati podatke, možemo se osloniti na veći broj sekundarnih servera za čitanje podataka. Replikacijom podataka na sekundarne servere ćemo se osigurati da u slučaju otkaza primarnih, možemo ih zameniti bilo kojim od sekundarnih.

## Predlog strategije za keširanje podataka

Najčešći upiti će se izvršavati nad klinikama i doktorima i možemo iskoristiti njihove ocene kako bismo odabrali najpopularnije i njih keširali radi bržeg pristupa. U slučaju da neka klinika ili doktor u poslednje vreme dobijaju puno pozitivnih ocena, možemo njima zameniti neki od keširanih klinika ili doktora koji u poslednje vreme ne dobijaju puno zahteva.

## Okvirna procena za hardverske resurse potrebne za skladištenje svih podataka u narednih 5 godina

Pretpostavimo da će najveći deo pristiglih zahteva za pregled biti zakazan i održan. Kako se zahtevi za pregled/operaciju brišu nakon što su zakazani, posmatračemo samo veličinu zakazanih pregleda/operacija i informacija koje se unose pri njihovom izvršavanju. Za zakazani pregled to je ukupno 44 bajta, a za završeni appointment se dodaje maksimum 280. Takođe se za svaki pregled može čuvati lista recepata koji su svaki po 40 bajtova. Računaćemo da za pregled ne bi trebalo da bude napisano preko 5 recepata.

Sabiranjem ovih veličina dobijamo  $44 + 280 + 5 * 40 = 524$  bajta po pregledu. Ako uzmemo kao pretpostavku cifru od 1 mil pregleda po mesecu, za 5 godina skladištenja će nam biti potrebno  $524 * 1 \text{ mil} * 12 * 5 = 31\,440\,000\,000$  bajtova, odnosno 29 TB.

Što se tiče registrovanih korisnika, ako pretpostavimo da su svih 200 miliona pacijenti, od kojih svaki zauzima maksimum 2100 bajtova, to nam daje cifru od 420 000 000 000 bajtova odnosno 391 TB.

Ukupno, krajnja procena za ukupno skladištenje je 420 TB za sledećih 5 godina. Iz računice su izostavljeni većina manje bitnih entiteta koji su relativno nepromenljivi kao i broj novoregistrovanih korisnika, za koji nemamo procenu.

## Predlog strategije za postavljanje load balansera

Load balanseri bi zasigurno bili postavljeni između krajnjih korisnika i aplikativnih servera radi smanjenja opterećenja servera, gde bi verovatno bili korišćeni neki od Least Connection ili Weighted Least Connection algoritama. Time bismo postigli da u slučaju da se neki od aplikativnih servera prepuni, na primer medicinskim osobljem koji su ulogovani u toku celog dana, ostatak saobraćaja rutira na server koji je manje opterećen.

## Predlog koje operacije korisnika treba nadgledati u cilju poboljšanja sistema

Delove sistema koje bi trebalo nadgledati:

- Za koje klinike pristiže najviše zahteva za pregled/operaciju
- Klinike i doktore sa najvećim brojem ocena
- Broj novoregistrovanih korisnika po danu
- Učestalost pristupa zdravstvenom kartonu

Kompletan crtež dizajna predložene arhitekture

