

Skalabilnost

U slučaju kada broj istovremenih korisnika preraste mogućnosti jednog server predlažemo sledeća rješenja:

- Prelazak sa monolitne na mikroservisnu arhitekturu
 - Mogućnost korištenja više servera
 - Skaliranje u realnom vremenu u zavisnosti od opterećenja sistema
 - Povećanje raspoloživosti servisa isporukom u više zona širom svijeta
 - Mogućnost balansiranja između raspoloživih servera radi smanjenja opterećenja servera i poboljšanja performansi
- Koristiti CDN za smještanje statičkih resursa
- Korištenje indeksa za pretragu (npr. Elasticsearch, Apache Lucene, Apache Solr i slično)
- Korištenje gRPC tehnologije umjesto REST-a
- Prelazak na noSQL bazu podataka
- Smanjiti količinu podataka za obradu i transport korištenjem paginacije
- Smanjiti opterećenje baze podataka keširanjem
- Smanjiti količinu protoka podataka keširanjem statičkih resursa i korištenjem HTTP 2 protokola
- Minimizirati JavaScript i CSS
- Izbaciti nepotrebne zavisnosti (nepotrebne biblioteke ili framework-e)

