



Primjer web aplikacije za pohranu i pregled prikupljenih mjernih podataka

Goran Horak
Velimir Gašparović

Calyx d.o.o.

Sadržaj

- Opis sustava
- Arhitektura sustava
 - Server
 - Klijent
 - Struktura podataka
- Deployment
- Post Mortem

Namjena sustava

- Unos, arhiva i pregled podataka
 - Podaci prikupljeni s distribuiranih (heterogenih) mjernih stanica
- Javno dostupan pregled agregiranih podataka
- Tri modula
 - Editor (unos, manipulacija, pregled, izvoz)
 - Admin (upravljanje korisnicima)
 - Javno sučelje (agregirani podaci, grafovi)

Editor sučelje

- Upravljanje općim podacima o mjernim stanicama
- Uvoz mjernih podataka (.txt, .csv, .xls, .xlsx)
- Pregled mjernih podataka
- Manipulacije mjernim podacima
 - Brisanje podataka
 - Promjene vremenskih zona
- Izvoz podataka

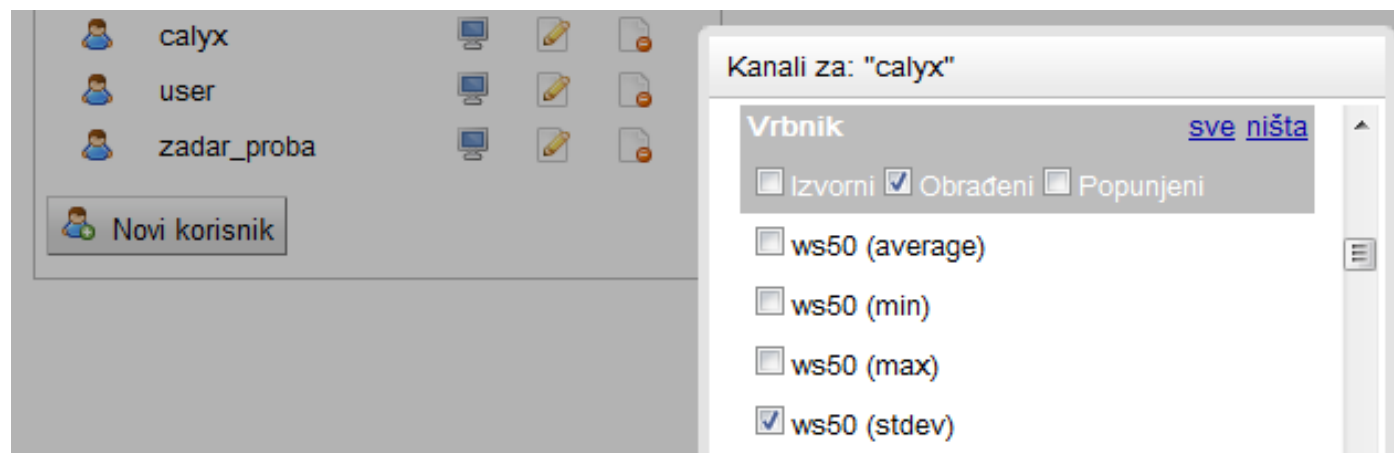
Podaci o mjernom stupu

Naziv:	Kasumi	Geografska širina:	
Opis:	awsercro W03	Geografska dužina:	
Komentar:		Nadmorska visina:	
Period između zapisa:	10	Oznaka koordinatnog sustava:	GK 5.zona

Naziv	Tip	Visina	Mjerni podaci	Dec.						
WS10	Anemometar	10	01.06.2007. - 31.07.2010.	2			Σ			
WS30	Anemometar	30	01.06.2007. - 31.07.2010.	2			Σ			
WS44	Anemometar	44	01.06.2007. - 31.07.2010.	2			Σ			
WS46	Anemometar	46	01.06.2007. - 31.07.2010.	2			Σ			
WD10	Vjetrulja	10	01.06.2007. - 31.07.2010.	1			Σ			

Administrativno sučelje

- Upravljanje korisnicima javnog sučelja
 - Prava na pojedine mjerne stanice
 - Prava na pojedina mjerenja s mjernih stanica
 - Vremenski okvir dostupnih podataka
 - Omogućavanje izvoza podataka



Javno sučelje

- Opći podaci o mjernim stanicama
- Prikaz agregiranih podataka
 - Srednja vrijednost, min-max vrijednost, dostupnost...
- Grafički prikaz mjernih podataka

Pregled podataka

Mjerna stanica:

Inačica:

Razdoblje:
od:

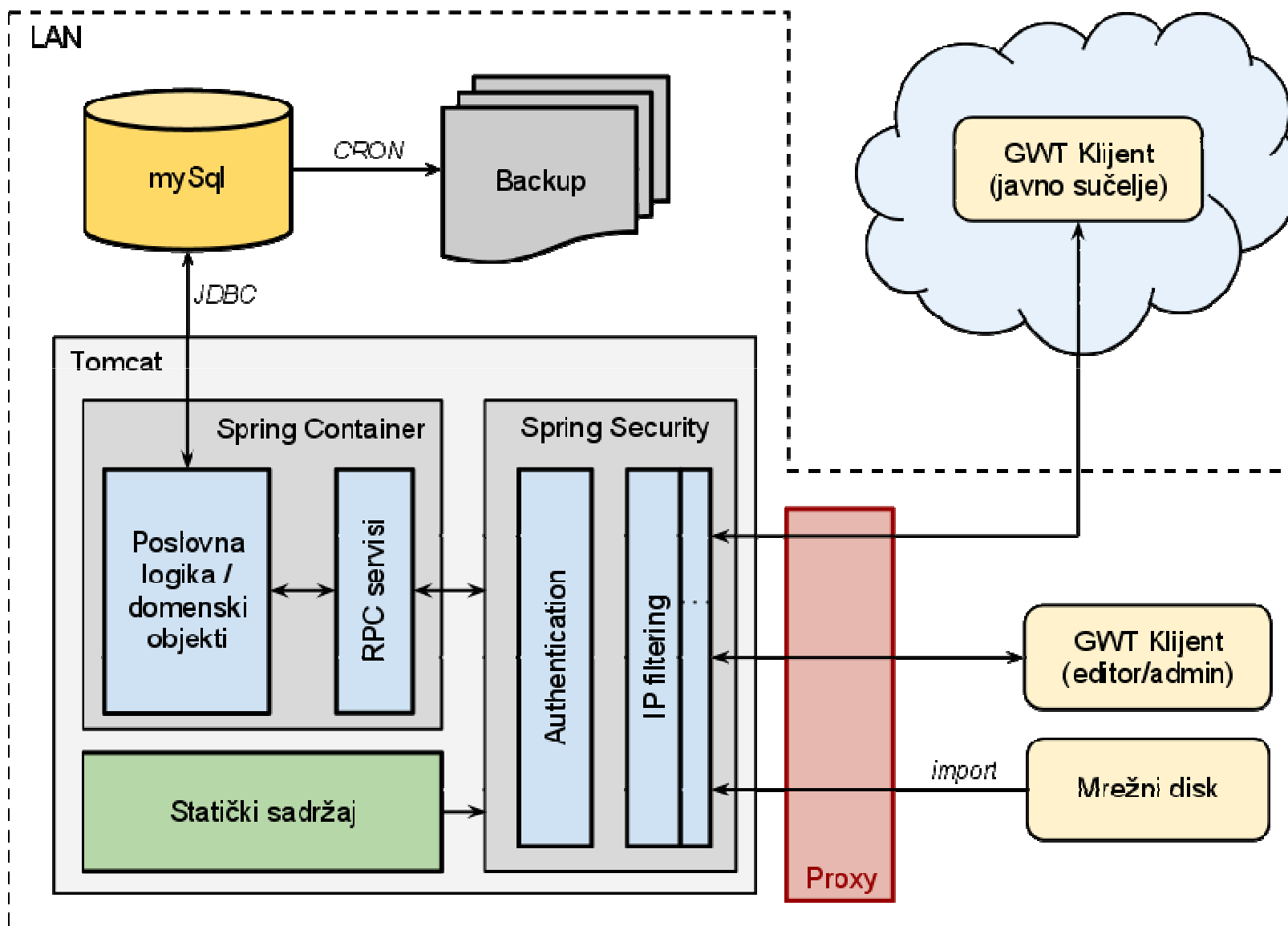
do:



Arhitektura sustava

- Server
 - mySql baza podataka
 - Tomcat servlet container
 - Java backend
 - Spring Framework
- Klijent
 - Google Web Toolkit

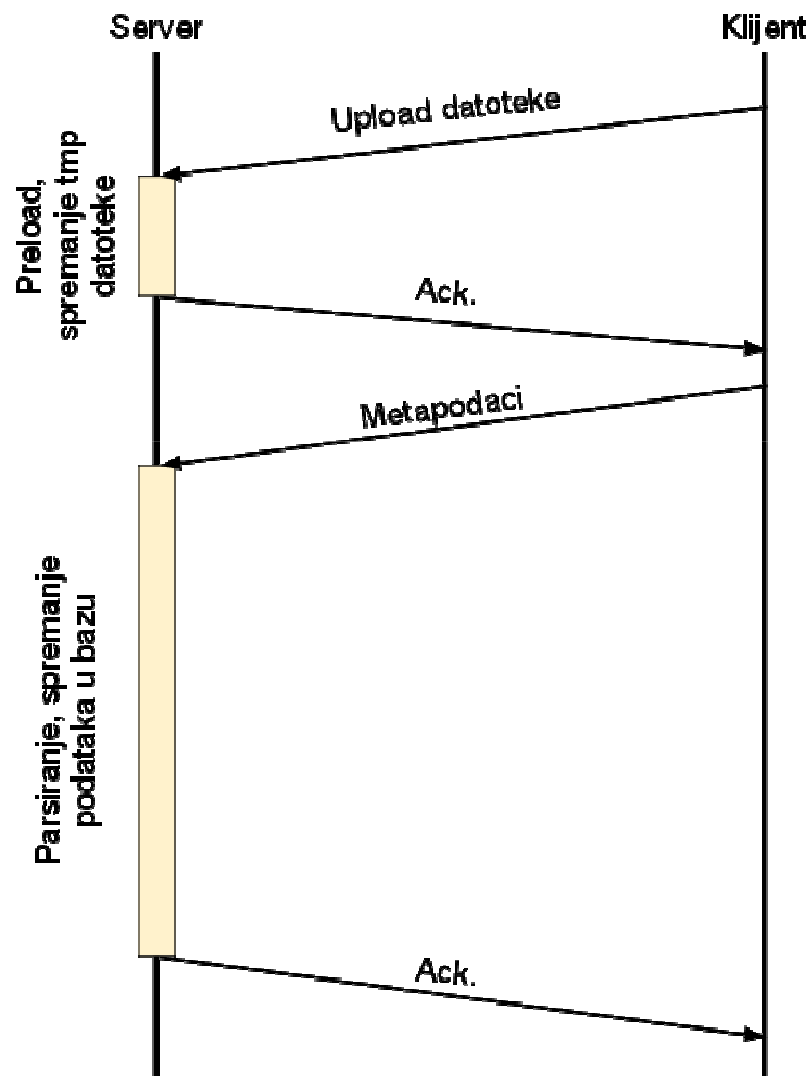
Arhitektura sustava



Klijent/server Komunikacija

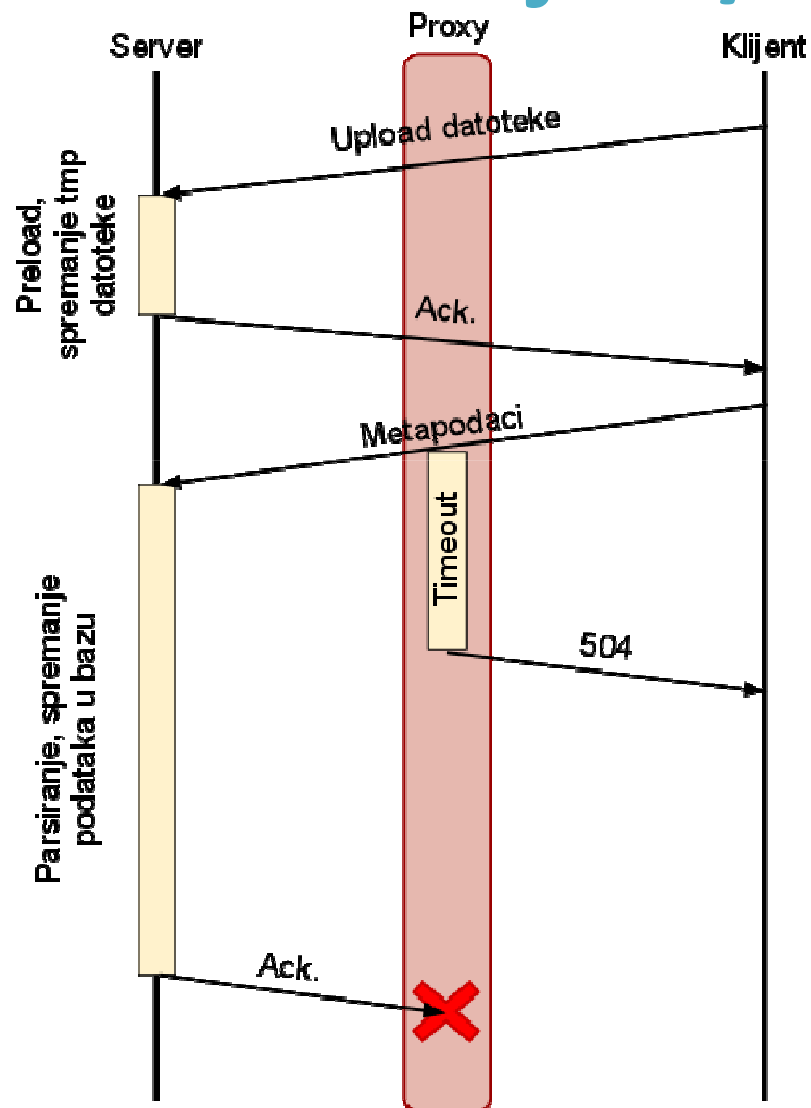
- GWT klijent
 - HTML/Javascript stranica
 - Asinkrona RPC komunikacija sa serverom
 - Uvoz podataka
 - Podaci se uvoze iz velikih datoteka
 - Dugotrajan proces (~ 2min)
 - Problem – proxy timeout

Klijent/server kom.



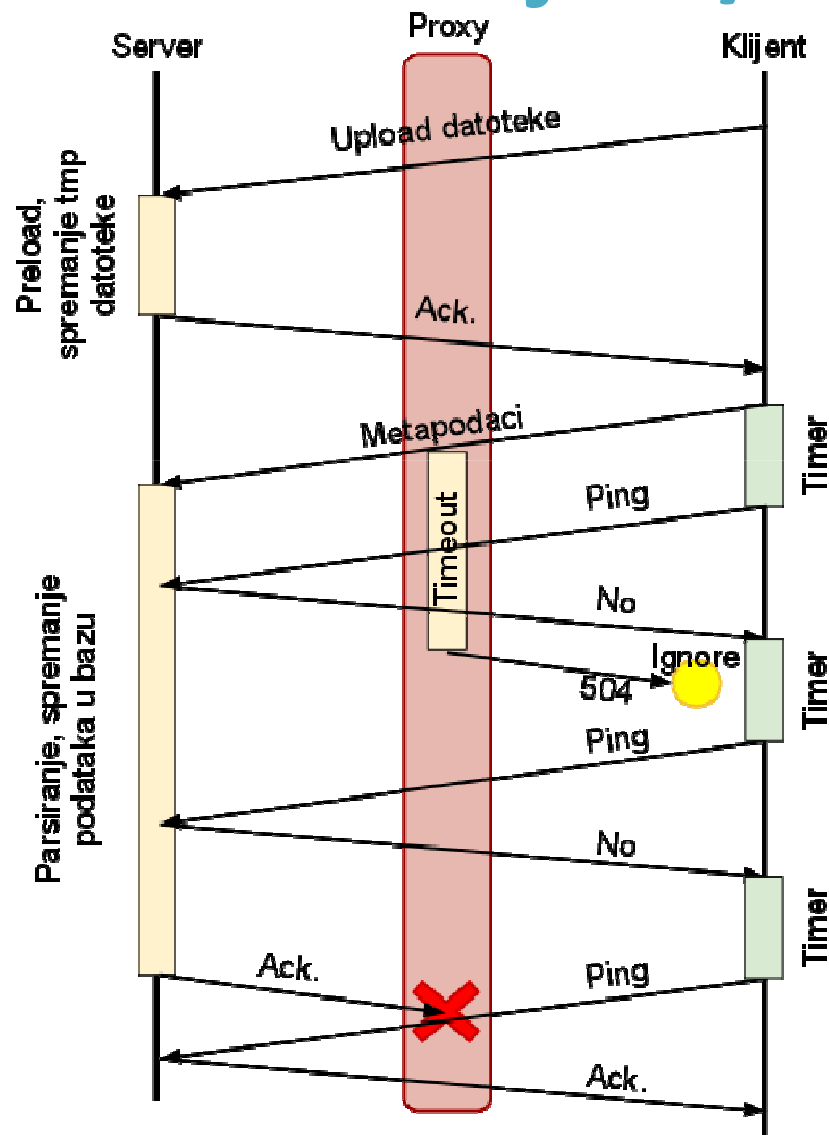
- Komunikacija bez proxy-ja
- Beskonačan timeout

Klijent/server kom.



- Proxy s određenim (prekratkim) timeoutom
- Nemogućnost push-a

Klijent/server kom.



- Uvođenje periodičkog provjeravanja servera sa strane klijenta

Struktura (baze) podataka





- Dva tipa podataka
 - Skalarni podaci s različitim mjernih stanica
 - Uzorkovanje svakih 10 minuta
 - Rasteri podataka
 - Raster 300 x 300
 - Uzorkovanje svaka 3 mjeseca
- Povijest podataka – 10 godina

Struktura (baze) podataka

- Metapodaci – klasična 1-n struktura
 - Mjerno mjesto – mjerna funkcija – senzor – vrijeme
- Spremanje samih podataka – izazov
 - Svaka pogrešna procjena
 - Nezadovoljavajuće performanse
 - Min. 48 sati popunjavanja baze da se omogući testiranje

Struktura (baze) podataka

- Bez umjetnog primarnog ključa
- Minimizirani tipovi podataka
- Indeksiranje
- Horizontalno particioniranje

TABLE 'data'		
	<i>time_id</i>	SMALLINT
	<i>point_id</i>	MEDIUMINT
	<i>modality_id</i>	TINYINT
	<i>level_id</i>	TINYINT
	value	FLOAT
INDEX (time_id, point_id, modality_id, level_id)		
INDEX (point_id)		
INDEX (modality_id)		
INDEX (level_id)		
PARTITION BY HASH (time_id) PARTITIONS 256		

Deployment

- Djelomičan redizajn sustava – proxy problemi
- Inicijalni import podataka
 - 13 GB komprimiranih binarnih podataka
 - 700 GB mySQL data file
 - 8 dana
 - Proveden u kući
 - Prenošnje mySQL data file-a kod klijenta

Post mortem

- Naučili smo (barem) dvije stvari
 - Proxy problemi – nismo dovoljno dobro proučili ciljanu okolinu i detaljno sastavili sistemske zahtjeve
 - Količina podataka – pouzdali smo se u osobnu “procjenu”, nismo formalno analizirali utjecaj stvarne količine podataka na performanse sustava na vrijeme

Pitanja i komentari

- ...