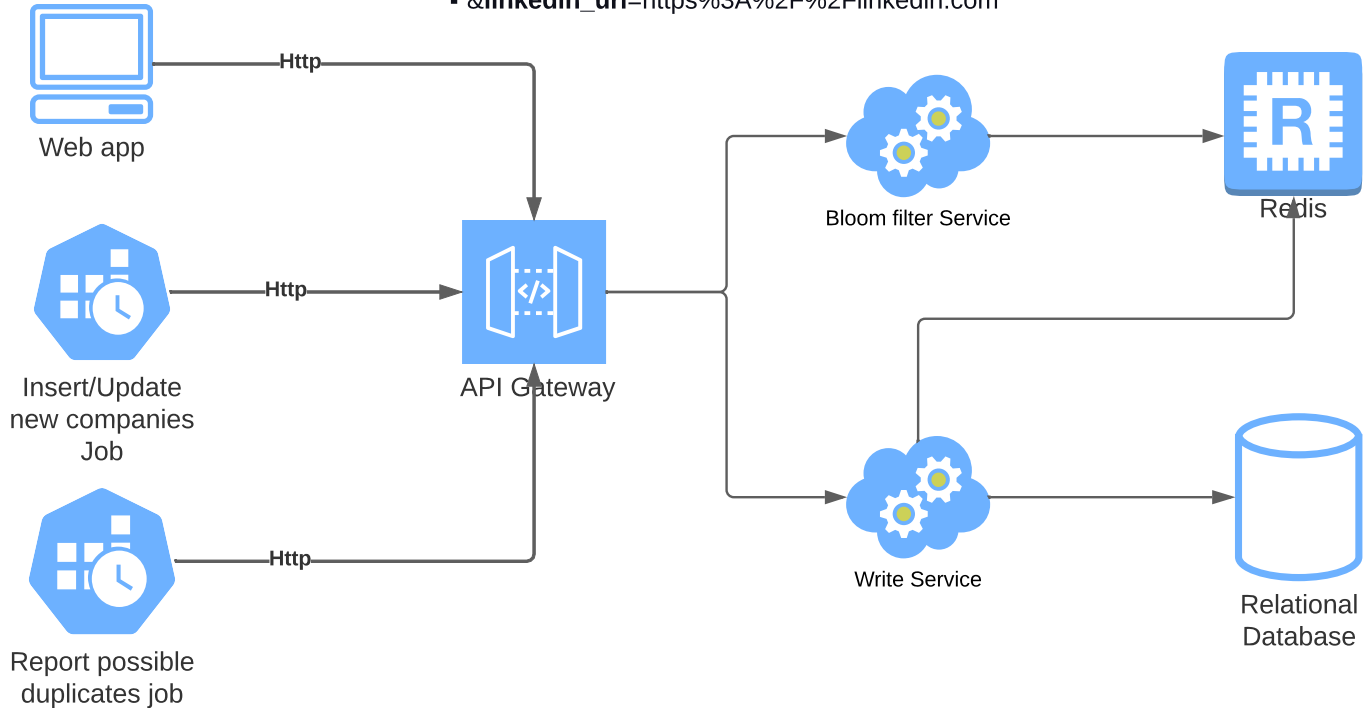API Endpoints
- PATCH /api/companies/:companyId
- POST /api/companies
- GET /api/companies
  - ?**name**=company1
  - &**short_name**=c1
  - &**website_url**=https%3A%2F%2Fgoogle.com
  - &**linkedin_url**=https%3A%2F%2Flinkedin.com

Web app

Insert/Update
new companies
Job

Report possible
duplicates job

Http

Http

Http

API Gateway

Bloom filter Service

Redis

Write Service

Relational
Database

- I have considered switching the database solution to a NoSql database but, as consistency tends to be important for company information, the relational database sounds like the best solution;
- One of the tradeoffs of the Redis cache would be data synchronization between the cache and the relational database;
- The "report possible duplicates job" would be executed by a scheduler and would generate a report with the possible duplicates;
- The Bloom filter is probabilistic so false positivies are expected.