

Analysing Big Data

Analysing Big Data

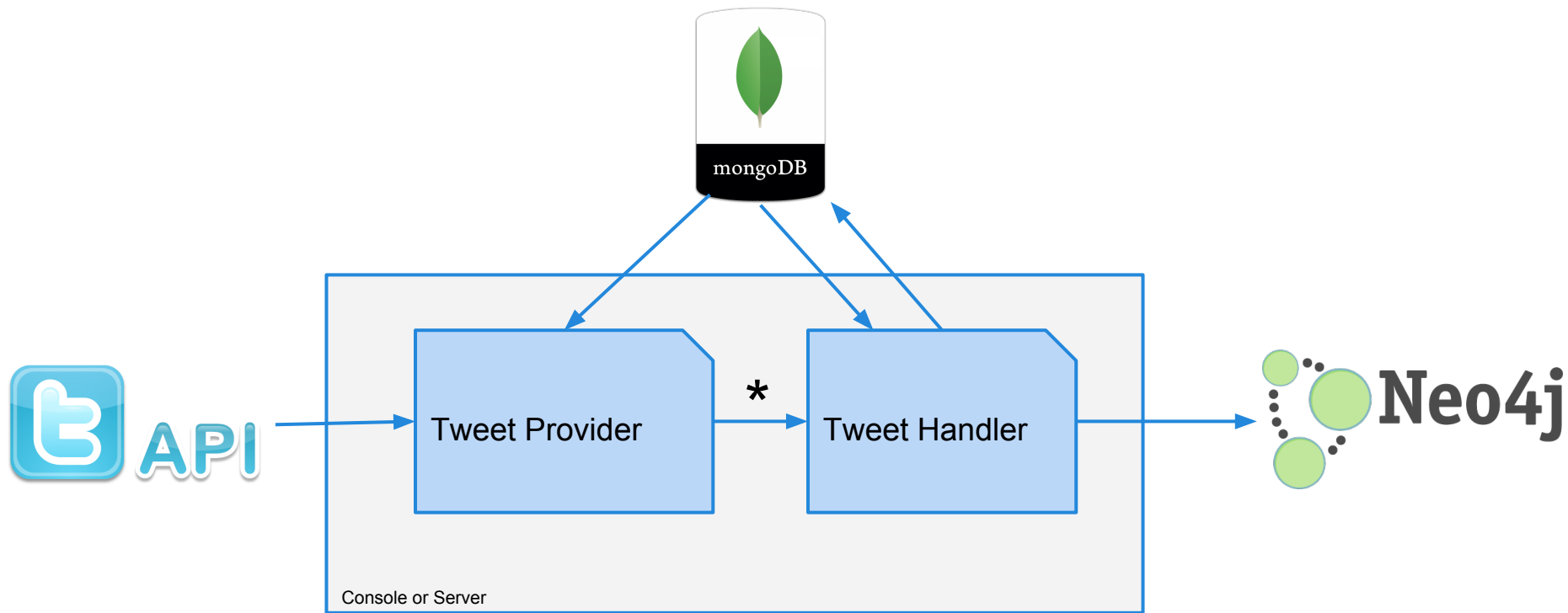
- Goal:
 - Find User interests
 - Advertise to them
- Basics:
 - Tweets
 - User Relationships
 - Homophily



Tasks

- create architecture
- mine tweets
- find topic interests
- find user relationships

Architecture



Tweet mining

- RestAPI vs. StreamingAPI
- Starting users?
- Watch users over days
 - Dump to mongoDB
- Search for specific terms
- Realised as Tweet Provider

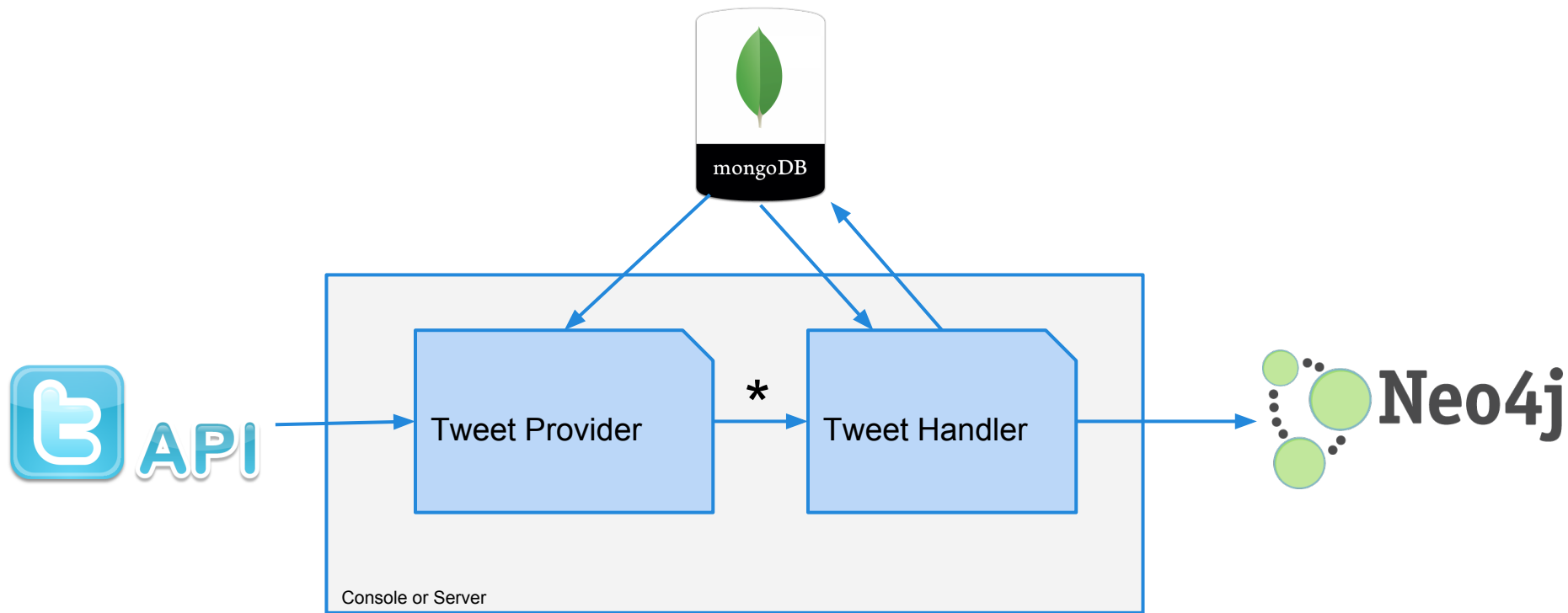
Find topic interests

- MongoDB Collections
 - one with topics
 - one with ads
- Topic mentioned twice -> interest -> relationship added in Neo4J DB between Topic and User
- Realised as Tweet Handler

Find User Relationships

- RestAPI limiting
- Use retweet information
- Store retweets
- Weighted edge in neo4j
- Realised as Tweet Handler

Architecture



Demo

Thanks