Large-Scale and Multi-Structured Databases

LondonSafeTravel

Ricky Marinsalda

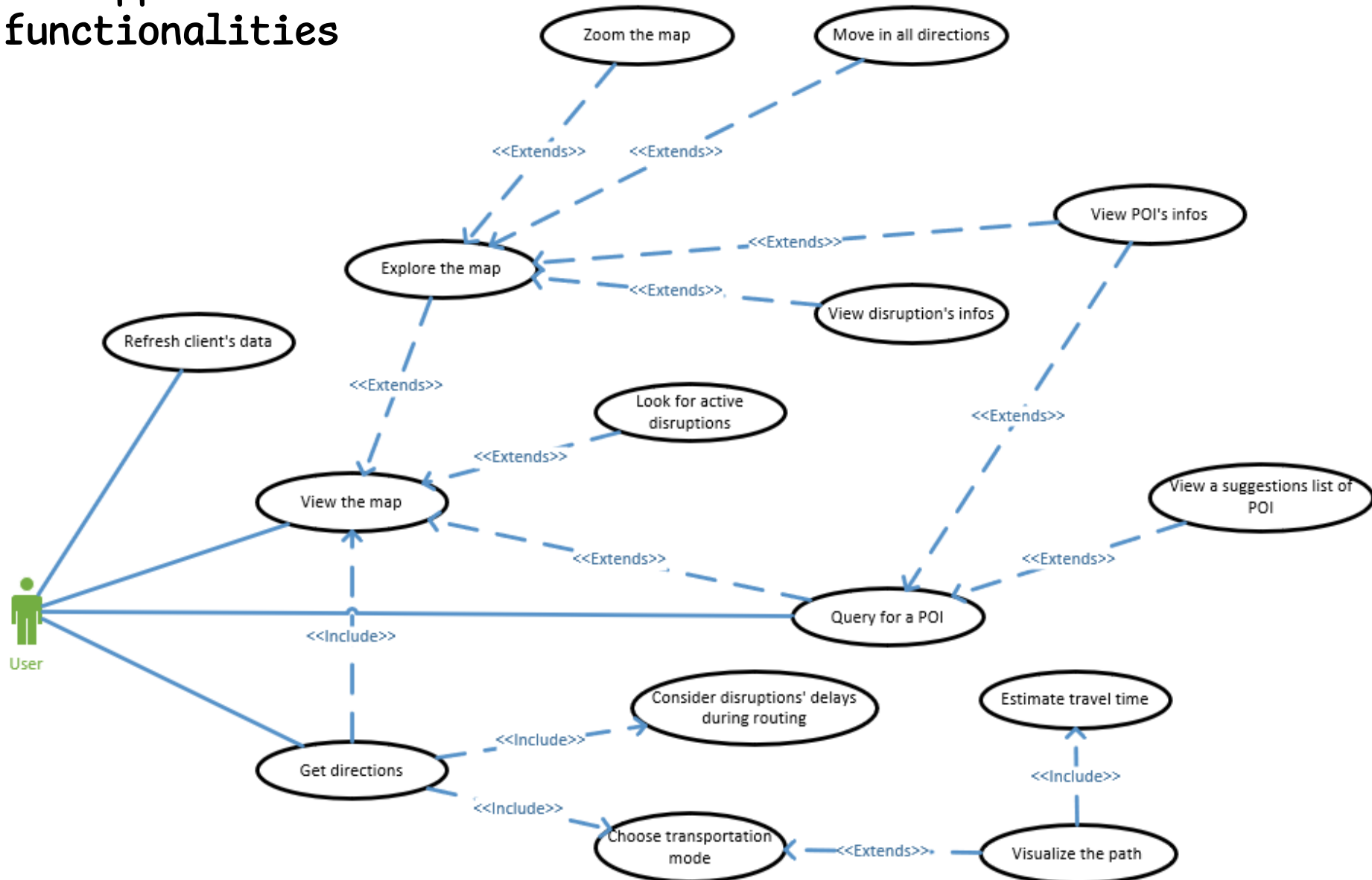Federico Frati

Dario Pagani

DIPARTIMENTO DI
INGEGNERIA
DELL'INFORMAZIONE

IN SUPREMÆ DIGNITATIS
1343

UNIVERSITÀ DI PISA

CROSSLAB
Innovation for industry 4.0

Application Highlights
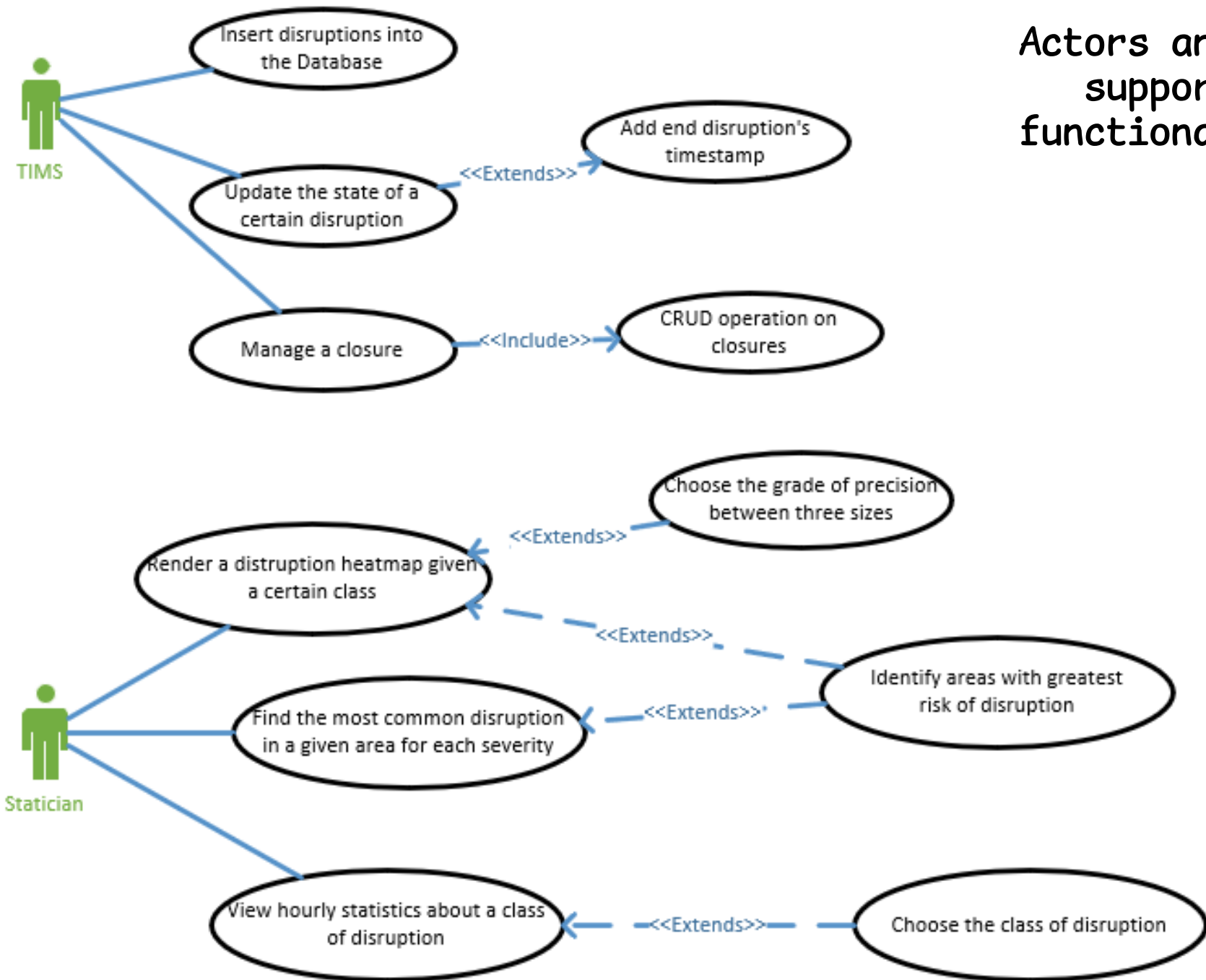
Our on-line service allows its users to explore, search and visualize a geographical map of London.

➔ Search for *POI*s and view their data
➔ Compute a travel route between two points
   · On foot
   · By car
   · By bicycle
➔ View and avoid disruptions on the network
   · Vehicles' accidents
   · Road works and other road closures

Actors and main supported functionalities

Actors and main supported functionalities

Insert disruptions into the Database

Update the state of a certain disruption

<<Extends>>

Add end disruption's timestamp

TIMS

Manage a closure

<<Include>>

CRUD operation on closures

Choose the grade of precision between three sizes

<<Extends>>

Render a distruption heatmap given a certain class

<<Extends>>

Identify areas with greatest risk of disruption

Statician

Find the most common disruption in a given area for each severity

<<Extends>>

View hourly statistics about a class of disruption

<<Extends>>

Choose the class of disruption

Dataset Description

Sources and Description:
>*OpenStreetMap* for the road network
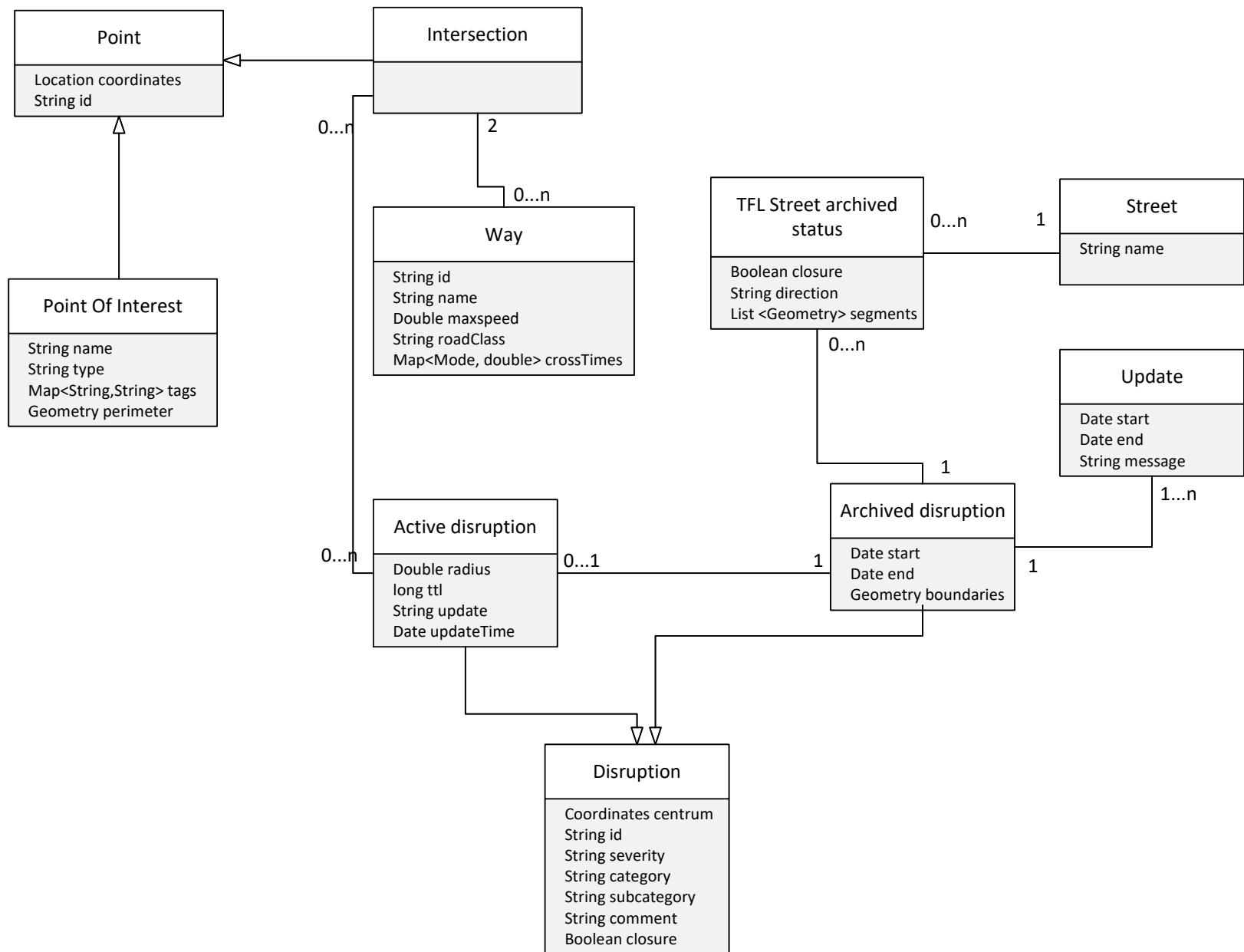>*Transport for London's TIMS API* for disruptions

Volume: circa 200MiB

Variety:
>*Multiple sources*
>*Multiple formats*

Velocity/Variability: *TIMS* data are like a live feed, older data are moved into an archive for future analytics

Point

Location coordinates
String id

Intersection

Way

String id
String name
Double maxspeed
String roadClass
Map<Mode, double> crossTimes

Point Of Interest

String name
String type
Map<String,String> tags
Geometry perimeter

TFL Street archived status

Boolean closure
String direction
List <Geometry> segments

Street

String name

Update

Date start
Date end
String message

Active disruption

Double radius
long ttl
String update
Date updateTime

Archived disruption

Date start
Date end
Geometry boundaries

Disruption

Coordinates centrum
String id
String severity
String category
String subcategory
String comment
Boolean closure

0...n   2   0...n   0...n   1   0...n   1   1...n   0...1   1   1   0...n

Requirements and Entities handled by Document DB

Entities

➤ *POI*s (attractions, landmarks etc..)

➤ Disruptions and their boundaries

Queries

➤ Search for *POI*s by name

➤ List *POI*s in a given area *[map view]*

➤ Number of active disruptions by class and hour of day

➤ Build a heat-map of a certain class of disruption

  ➤ *for broken traffic lights, vehicles collisions, burst water pipes, etc..*

➤ Most common disruptions by severity in a given area

Requirements and Entities handled by Graph DB
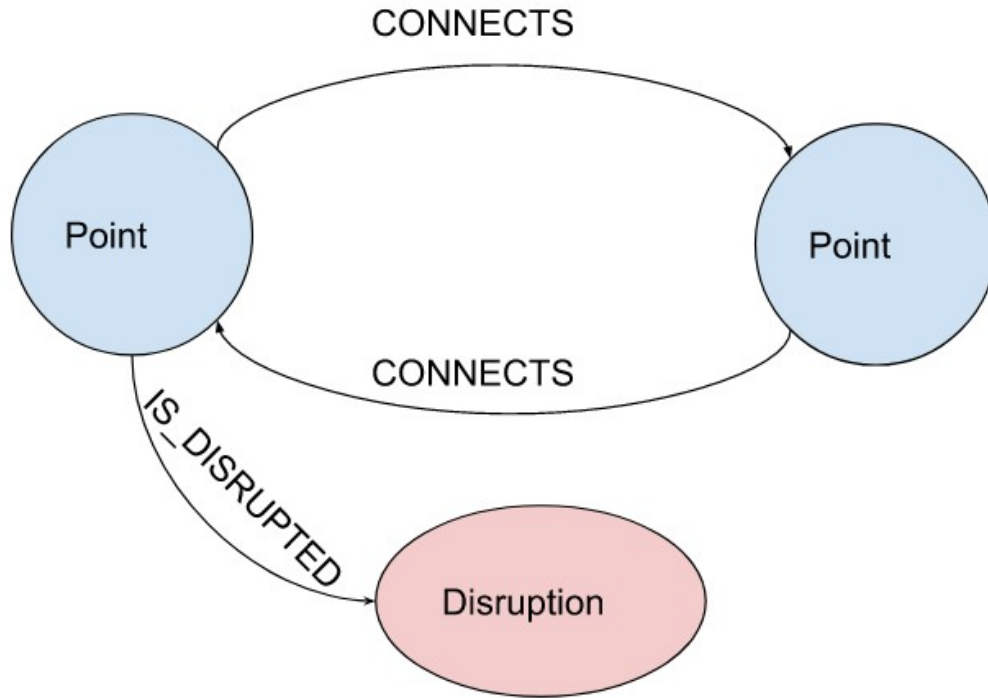
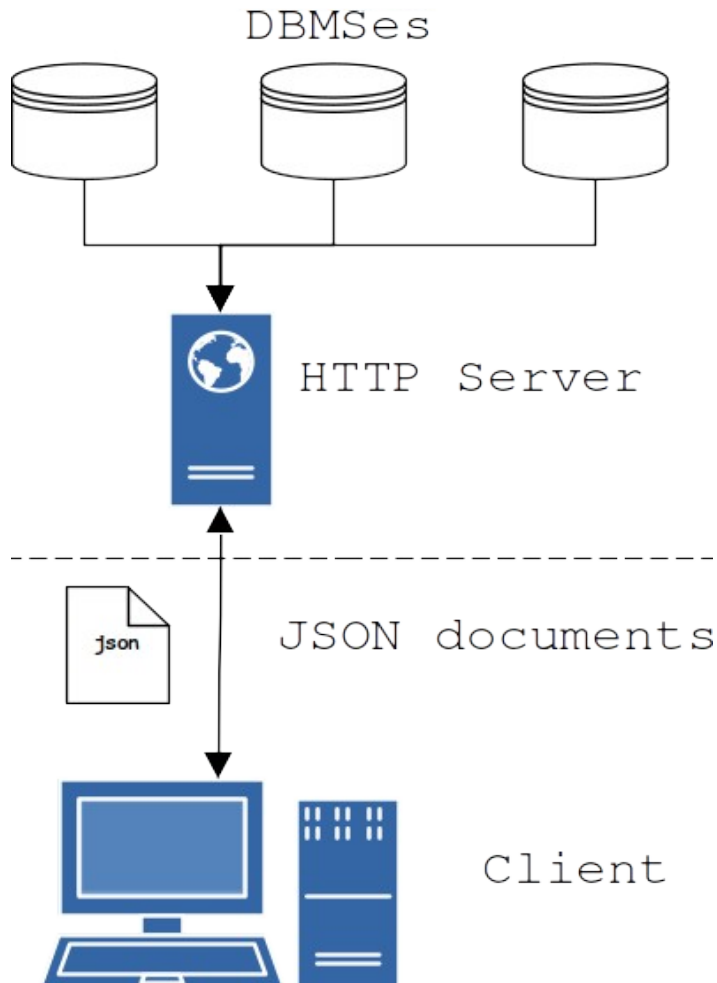**Query:**
Find a route between two points using AA*

Query:
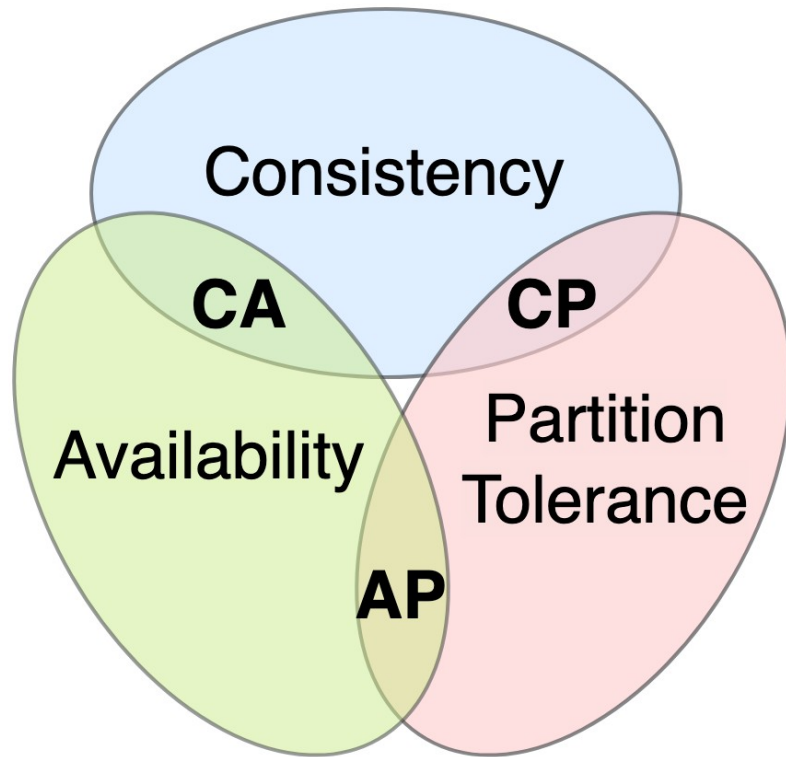Find and link all *Points* affected by a *Disruption* given its radius

Query:
Find a *reasonably* reachable *Point* given a mode of transportation

Software Architecture Preliminary Idea
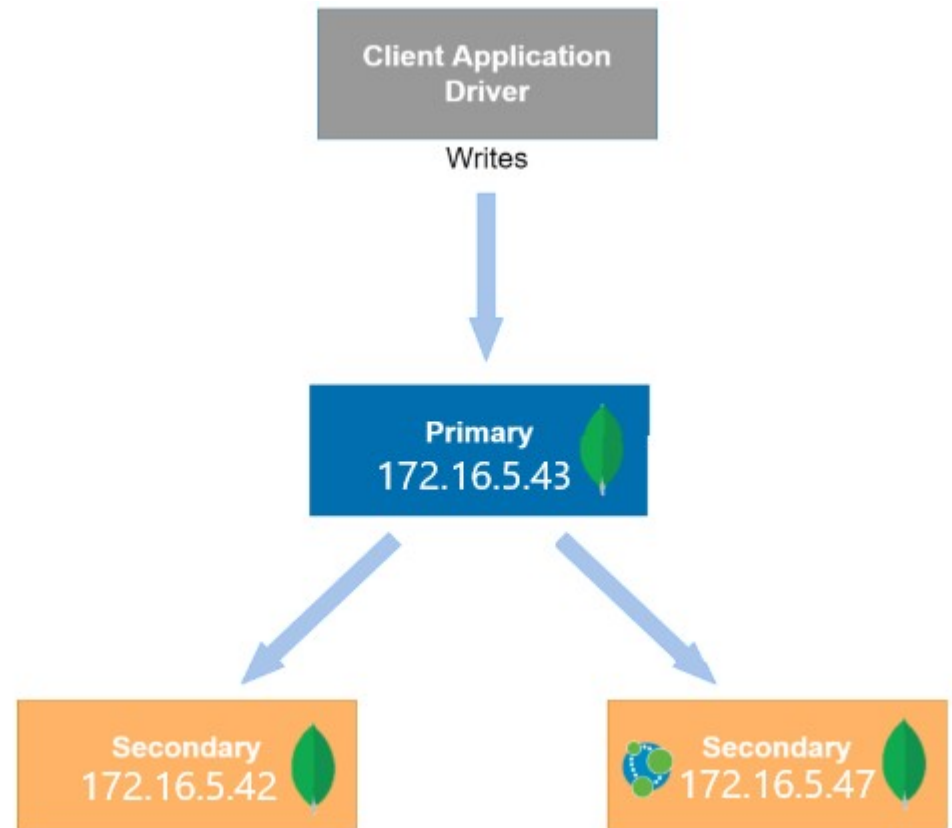
CAP thorem issue
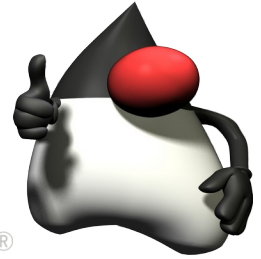

APPLICATION MODEL: **AP**

We prioritize:
   Availability
   Partition tolerance

We couldn't have data consistency

Client Application

Java Swing
    jxMapViewer2
    jxFreeChart

REST API
    Java Web Server
    Gson

Databases
    neo4j driver
    mongodb driver