Working with geodata in Go

Andrew Minkin
Who am I

• TeamLead at maddevs.io
• 10 years of experience
• https://github.com/meshbird/meshbird
Bishkek

• 1kk citizens

• 100++ taxi services
Namba Taxi

• 8k orders
• 600+ online drivers
• 500k clients
What’s a taxi

• Clients
  • Voice
  • Text
  • Mobile App
  • ....

• Drivers

• Operators

• Managers
AVG response time

• Drivers 20 ms
• Operators 2.5 ms
Prehistory
Машина в пути
First challenge

• Decrease update interval on driver’s device
  • Interval 15 secs
First try

• Make request -> save coordinates
• Make request -> animate the car
First problems

• We can’t animate car properly
• Car moves through fields, forests and quarters
First problems
Solution

• OSRM
Try again

• Timeout – 15 secs
• Make request -> save coordinates
• Send coordinates
• Building route via OSRM
• Get route -> animate marker
Машина в пути
Solution

• Check for 20 meters
Application Released!
Needs improvement

- Trip cost calculations on driver side
- We need more tracks and 1 track at 15 seconds is few
- GPS problems on driver
GPS problems

• Bad device/Bad module
• GPS going to die during the time
• Pits and “Jams”
Tasks

• Start to collect more tracks from drivers
• Show animated cars on the main screen
• Store intermediate trip cost on the server side
• Save mobile data
• Collect each track per one second
What’s the track?

- Latitude
- Longitude
- Session
- OrderID
- TripCost
Trafic economy

• 1 track = 100 bytes
Tell me all the options, please

- HTTP
- WebSockets
- TCP
- UDP
Tell me all the options, please

- HTTP
- WebSockets
- TCP
- UDP
Why we chose UDP?

• We send only datagrams
• We don’ t need guarantees
• Minimalism
• Save lots of data
• We have only 20 bytes overhead
• Not blocked in our country
What about data serialization?

• JSON 😞
• MsgPack 😞
• Protocol Buffers 😊
Data size

<table>
<thead>
<tr>
<th>Format</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protobuf</td>
<td>42</td>
</tr>
<tr>
<td>MsgPack</td>
<td>127</td>
</tr>
<tr>
<td>JSON</td>
<td>137</td>
</tr>
</tbody>
</table>
Total

- 42 bytes of payload
- + 20 bytes of IP headers
- = 62 bytes per track
- = PROFIT!
Data storage
What data to store?

• Driver’s session
• Cab number
• Order ID
• Trip cost
• Last location
• N last locations
Which storages do we use?

- Percona
- Redis
- Elasticsearch
We need geoindex

• KD-tree
• R-tree
What requirements for geoiindex

• Search of N nearest points
• Balanced tree
KD-Tree
KD-tree cons

- Unbalanced tree
- Can search only one nearest point
R-tree
Where can I get it?

• https://github.com/dhconnelly/rtreengo
What’s wrong with drivers?

• Bad internet connectivity
• Turned off the phone
• Low battery
• Removed application from RAM
• And lots of other reasons
• We need expire mechanism
• We need LRU data structure for storing coordinates
Storage architecture

• In-memory
• R-tree
• Map with drivers(key is the session)
• Map with drivers(key is cab number)
What algorithm on backend

- Get data by UDP
- Try to get driver from storage
  - If doesn’t exist – get driver from redis
- Check and validate data
- Set driver to storage
  - If doesn’t exist – initialize LRU
- Update r-tree
Go

• Strong typed and compiled
• Small size of docker containers
• Few resource usage
• Less own zoo 😊
HTTP API

• Return nearest drivers
• Remove driver from storage (by cab number or session)
• Get information about trip
• Get information about driver
How to maintain

- Logging into stdout/stderr
- Metrics to Graphite
- Checks to sensu
- Usefull /status
- Bots
Usefull /status

• Uptime since
• HTTP Statuses counters
• Total requests
Bots

• Emulate of driver’s workflow
• Emulate of client’s workflow
• Running near Marokko or Congo
How it looks now?

• We get client location from sensors
• We get nearest drivers with routes
• Animate each car
• Update interval 15 seconds
Main slide of first story

• UDP+Protobuf for data savings
• In-memory storage
• R-tree for nearest drivers
• LRU cache for storing last locations
• OSRM for map matching and building routes
Challenges in geocoding
What challenges do we have?

• Lack of data
• Crossroads navigation
• We can’t trust GPS
Which map providers available?

• Yandex
• Google
• 2GIS
What we have to solve issue

• OpenStreetMap data
• Own database with addresses with coordinates
Search format example

- Chui ave 139
- Chui / Manasa
- Red centre
- AUCA
- 5-15
- 7 marksa 181
We need own geocoder

• Available to search in different formats
• Support for auto translated values
• Support of synonyms
Elasticsearch index

• Addresses(alias)
  • OSM_timestamp
  • Drivers_data
./ariadna update

• Download file
• Create index in Elasticsearch
• Populate with data from osm
• Search intersections and populate data
• Change aliases
• Removes old index
Features

• Geocode
  • By the name
  • By crossroad
  • By the name of institution
  • Synonyms support
• Reverse geocode
Where to get it?

• https://github.com/maddevsio/ariadna
Questions?

• https://github.com/maddevsio
• https://github.com/maddevsio/ariadna
• @gen1us2k (twitter,github,facebook,skype,telegram)