



# Scalable cloud resources for improved cycling conditions

Author: Dr.-Ing. Philipp Grubitzsch

# ABOUT MOVEBIS - RESEARCH PROJECT FOR THE OPTIMISATION OF TRAFFIC CONDITIONS

[Movebis](#) is a research project funded by the mFund funding programme of the Federal Ministry of Transport and Digital Infrastructure from 2017 to 2020. Besides the TU Dresden with the involved chairs for transport ecology and computer networks, Cyface GmbH as subcontractor and Climate Alliance with several own subcontractors are also part of the project consortium. The aim of the project is to automatically determine information about the cycling infrastructure from sensor-based movement data of cyclists. The data on which the analysis is based will be collected during the annual [CITY CYCLING campaign](#) of Climate Alliance.

## ABOUT CITY CYCLING

CITY CYCLING is a competition in which the participants' aim is to cycle as many everyday routes as possible in a period of 21 days. On the one hand, this is intended to make the advantages of cycling practically tangible - especially for road users who would otherwise cover their distances by car. On the other hand, the campaign reaches out to local politicians who decide on cycling infrastructure measures for their communities. The smartphone app (iOS & Android) developed for CITY CYCLING provides the Movebis research project with anonymised sensor data of GPS, acceleration sensor, magnetometer and gyroscope. The resulting infrastructure information can be used to determine, for example, where, how much and how fast people cycle or where cyclists spend a particularly long time at traffic lights. This information will be processed in form of maps and statistics and will be made available in particular to planners and bicycle traffic representatives in local authorities.

## DIGITAL INFRASTRUCTURE – TECHNICAL REQUIREMENTS

In this process, sensor data from potentially hundreds of thousands of smartphones is collected and stored. A single participant alone can record many dozens of his trips. In 2019, around one million journeys were recorded by more than 77,000 participants.

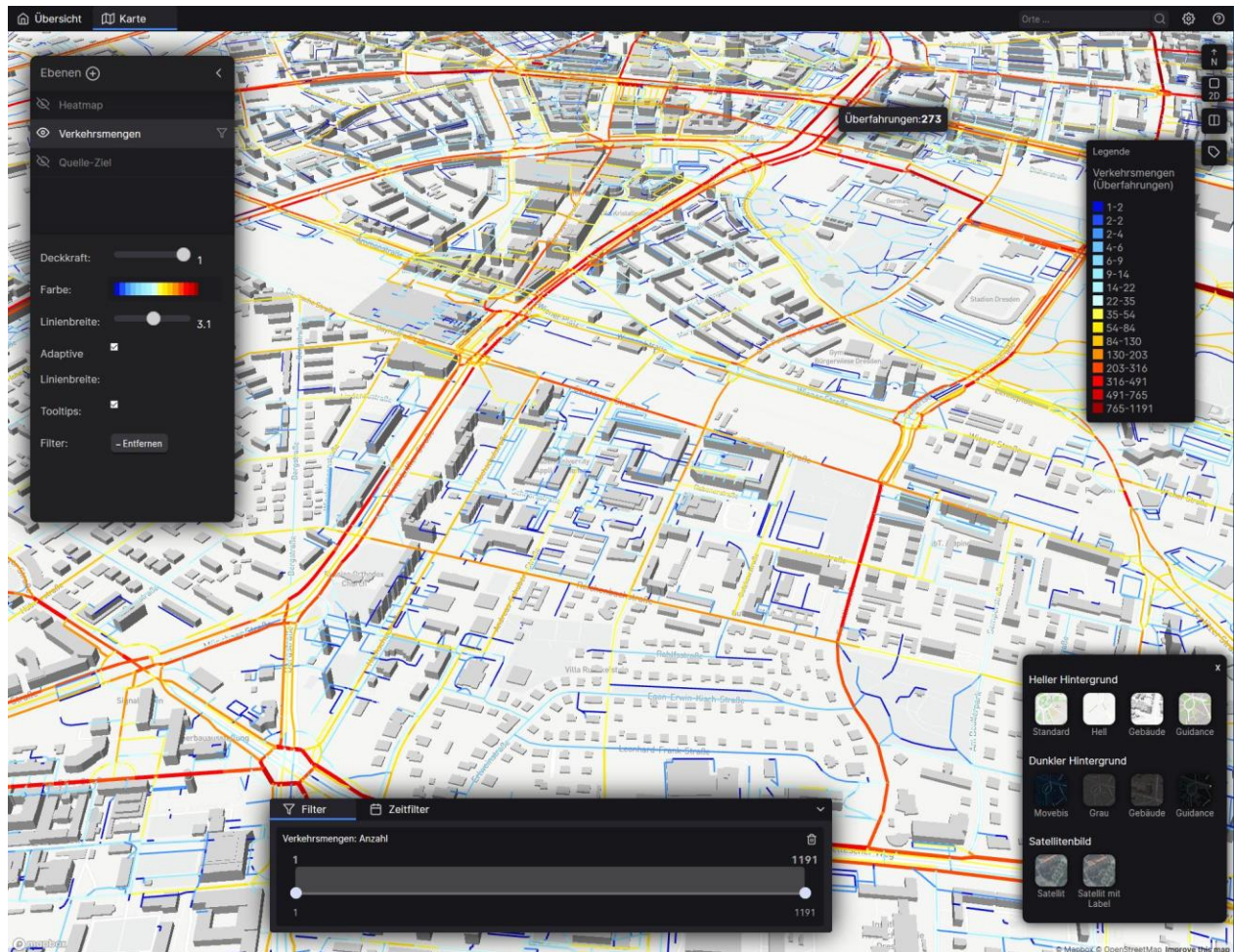


**Picture 1.** Data collection is carried out within the framework of the CITY CYCLING campaign using a special app, source: Climate Alliance

Since other sensors in addition to GPS will also record data, the amount of data per recording will vary from a few to several dozen megabytes - depending on how long the recording of a trip takes. In 2019, the total data volume was 5 terabytes.

One requirement is to convert the data into current results in real time. In order to achieve this with annually increasing user numbers, the consortium around Movebis decided on the cloud as a horizontally scalable, digital infrastructure solution. The data processing should also be highly available and be able to continue working without errors in the event of a node failure (e.g. a VM or server). Since 2019, Movebis

has also been working on the redundant data processing and storing and on making the delivery of the results high-performant and highly available for potentially several thousand of users.



**Picture 2:** Representation of the main traffic flows caused by bicycle traffic - the example of Dresden, source: TU Dresden

## SETUP OF INFRASTRUCTURE

The heart of the data processing is a distributed data stream processing engine on which use cases are developed that convert raw data into current results in real time. In addition, there are various services such as web services, message queues and load balancers to collect and distribute the data in the infrastructure, as well as distributed data storages and databases to store raw, status and result data. Other web services are needed for rendering geographical views of traffic data, delivering all results to users, for authentication purposes, etc. In all these contexts, there are few additional isolated microservices that perform specific tasks to enable the interaction of all distributed system components. All services in this research prototype are designed at least double or triple redundant. This extensive

service conglomerate is supported by containerization and orchestration methods. An additional web service and a time series database are necessary for monitoring the entire deployment.

In 2018, the infrastructure was only used with 4 VMs of very small and medium-sized flavours and 4 TB object storage. One year later, there were already 15 VMs of almost all flavours including several hundred gigabyte volume storage and 5 TB object storage.

The Cloud&Heat dashboard is used to manage the basic configuration of the deployment environment. Here, the VMs of different flavours required for data processing are managed and configured with the operating system we require. The configuration of VMs in different availability zones is also important for the deployment, otherwise it would not be possible to realize a reasonable high availability operation. Several replication instances of a service must necessarily be located on different parts of the hardware infrastructure. In addition to the local SSD storage of a VM, the mounted volume storage is configured for the VMs. The object storage is mainly used for reliable intermediate storage of the received raw data. The local and external network configuration, firewall rules, floating IPs and key pairs for access are also very important for distributed deployment. All in all, up to 21 VM instances with a total of 78 vCPUs, 247 GB RAM, 2 TB SSD and volume storage in three availability zones will be in use in 2020.

## WHAT'S NEXT?

Based on the feedback from the pilot municipalities, but also from various public presentations, it is planned to transfer the results of the research project into a company from 2021 onwards, which will ensure the continued operation of the data evaluation. The further development of concrete use cases for the collection of the cycling infrastructure but also other mobility data in the public space can also be done by this company or will possibly be realised in further research projects based on Movebis. For the data processing a corresponding setup will be needed in the future as well. Cloud&Heat has proven to be a reliable service provider during the project period and could also establish itself as a permanent infrastructure provider in the context of a spin-off.

Semi-automated orchestration and dynamic scalability at runtime are important points for the further development of the deployment. Here, the Managed Kubernetes service offered by Cloud&Heat can serve as a basis. In combination with secustack's technology for securing the rolled out containers, a new level of security is achieved with regard to the sensitivity of the processed data.