# MovR as a Benchmark for Geo-distributed Databases

## Performance Evaluation and Insights

Wilhelm Marcu (w.p.a.marcu@student.tudelft.nl)

Responsible professor: Asterios Katsifodimos | Supervisors: Oto Mráz, George Christodoulou, Kyriakos Psarakis

## 1. Introduction

- Distributed systems: high scalability, fault tolerance
- Geo-distributed databases (GDDB)
- Low latency, high availability

- Problem: benchmarks not GDDB specific
- Goal: benchmark to properly test GDDB

## 2. Research Question

How effective is the MovR benchmark for evaluating the performance characteristics of GDDB in terms of throughput, latency, bytes transferred, cost?

Baseline 

Skew 

Scalability 

Network 

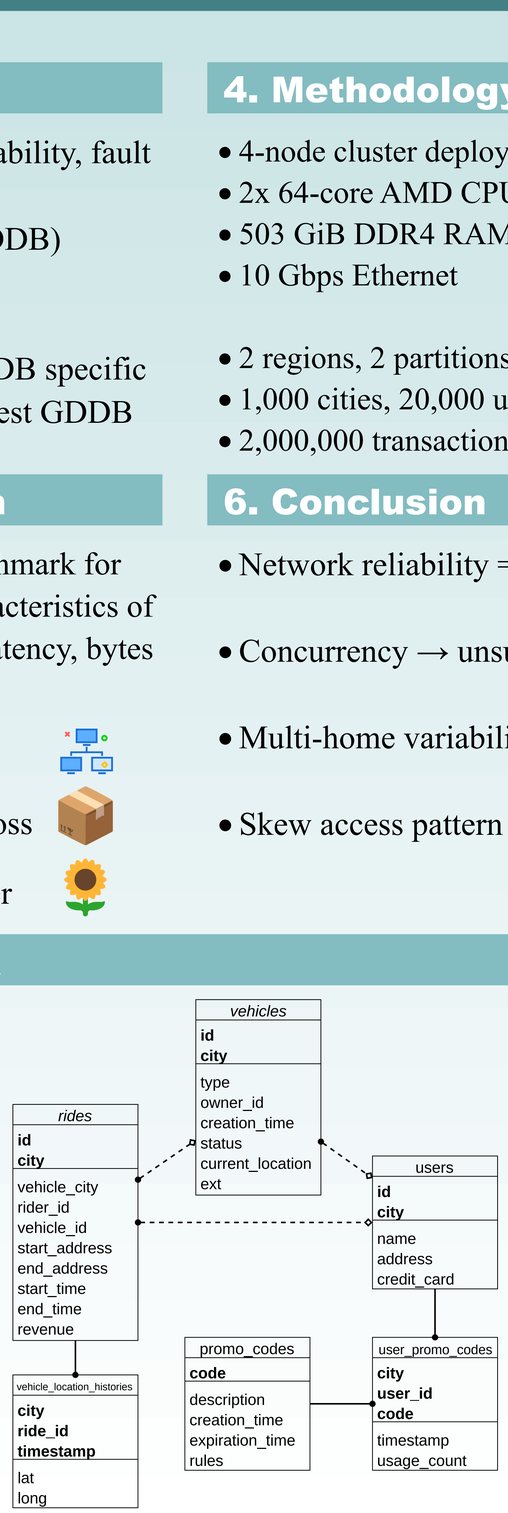
Packet Loss 

Sunflower 

## 3. Benchmark - MovR

- Vehicle sharing service
- CRUD + time-series
- Geographic data locality

1. Find ride: read
2. Add user: write
3. Add vehicle: write
4. Start ride: read/write/update
5. Track location: write
6. End ride: update




## 4. Methodology

- 4-node cluster deployment
- 2x 64-core AMD CPU
- 503 GiB DDR4 RAM
- 10 Gbps Ethernet

- 2 regions, 2 partitions
- 1,000 cities, 20,000 users, 5000 vehicles
- 2,000,000 transactions

## 6. Conclusion

- Network reliability = major bottleneck
- Concurrency → unsustainable latency
- Multi-home variability little impact
- Skew access pattern not strong enough

## 5. Results