



1. Introduction

- Distributed systems: high scalab tolerance
- Geo-distributed databases (GD]
- Low latency, high availability
- Problem: benchmarks not GDD
- Goal: benchmark to properly tes

2. Research Question

How effective is the MovR benchr evaluating the performance charac GDDB in terms of throughput, late transferred, cost?

Baseline

Skew



Network

Packet Los



Sunflower

3. Benchmark - MovR

- Vehicle sharing service
- CRUD + time-series
- Geographic data locality
- 1. Find ride: read

Scalability

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

	4. Methodology	5.
bility, fault	 4-node cluster deployment 2x 64-core AMD CPU 	
DB)	 • 503 GiB DDR4 RAM • 10 Gbps Ethernet 	(s/ux) 40000
OB specific est GDDB	 2 regions, 2 partitions 1,000 cities, 20,000 users, 5000 vehicles 2,000,000 transactions 	60000 Function (
	6. Conclusion	f) 40000 hdybri 20000
mark for cteristics of	 Network reliability = major bottleneck 	Thro
tency, bytes	• Concurrency \rightarrow unsustainable latency	ut (txn/s)
	 Multi-home variability little impact 	10 40000 10 40000 10 20000
SS I I I I I I I I I I I I I I I I I I	• Skew access pattern not strong enough	(x1/s) (x
	vehicles id city	
rides id city vehicle_city rider_id vehicle_id start_address end_address start_time	type owner_id creation_time status current_location ext id city name address credit_card	60000 40000 20000 (
end_time revenue vehicle_location_histories city ride_id timestamp lat long	promo_codes user_promo_codes code city description user_id creation_time code expiration_time timestamp rules usage_count	Lhroughput (txn/s) 30000 10000 (

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

Results

