Distributed Spatio-Temporal Database Management System

Mohamed Bakli, PhD student
Supervisor: Esteban Zimányi (ULB), Advisor: Mahmoud Sakr (ULB)

Existing Systems

MobilityDB

citusdata

MobilityDB

Q10:Which vehicles travelled within one of the regions from Regions1 during the periods from Periods1?

```
EXPLAIN SELECT DISTINCT R.RegionId, P.PeriodId, P.Period, C.Licence
FROM Trips T, Cars C, Regions R, Periods P
WHERE T.CarId = C.CarId
AND T.trip && STBox(R.geom, P.Period)
AND ST_Intersects(trajectory(atPeriod(T.Trip, P.Period)), R.geom)
ORDER BY R.RegionId, P.PeriodId, C.Licence;
```

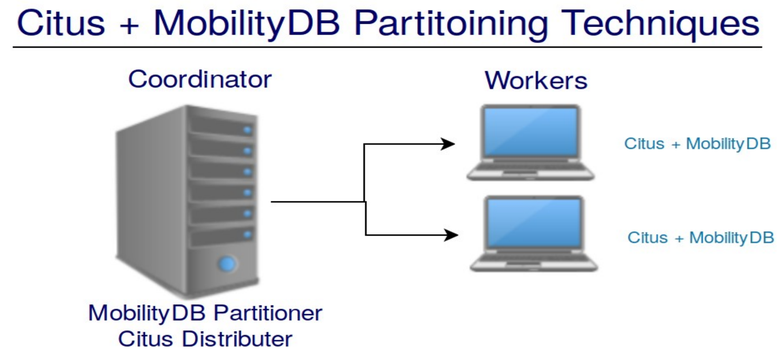
Single Node Query Plan:

```
Unique  (cost=528372.26..528372.87 rows=49 width=40)
  ->  Sort  (cost=528372.26..528372.38 rows=49 width=40)
        Sort Key: regions.regionid, p.periodid, c.licence, p.period
        ->  Nested Loop  (cost=195.16..528370.88 rows=49 width=40)
              ->  Nested Loop  (cost=194.88..528356.41 rows=49 width=36)
                    ->  Nested Loop  (cost=0.00..3.08 rows=100 width=875)
                          ->  Seq Scan on regions  (cost=0.00..13.00 rows=100 width=847)
                    ->  Materialize  (cost=0.00..0.45 rows=10 width=28)
                          ->  Subquery Scan on p  (cost=0.00..0.40 rows=10 width=28)
                                ->  Limit  (cost=0.00..0.30 rows=10 width=44)
                                      ->  Seq Scan on periods  (cost=0.00..3.00 rows=100 width=44)
              ->  Bitmap Heap Scan on trips t  (cost=194.88..5283.52 rows=1 width=36)
                    Recheck Cond: (trip && stbox(regions.geom, p.period))
                    Filter: ((trajectory(atperiod(trip, p.period)) && regions.geom) AND _st_intersects(trajectory(atperiod(trip, p.period)), regions.geom))
                    ->  Bitmap Index Scan on trips_spgist_idx  (cost=0.00..194.88 rows=1466 width=0)
                          Index Cond: (trip && stbox(regions.geom, p.period))
        ->  Index Scan using cars_carid_idx on cars c  (cost=0.28..0.30 rows=1 width=12)
              Index Cond: (carid = t.carid)
```

Multi-Node Cluster Query Plan:

```
Sort  (cost=0.00..0.00 rows=0 width=0)
  Sort Key: remote_scan.regionid, remote_scan.periodid, remote_scan.licence
  ->  HashAggregate  (cost=0.00..0.00 rows=0 width=0)
        Group Key: remote_scan.regionid, remote_scan.periodid, remote_scan.licence, remote_scan.period
        ->  Custom Scan (Citus Real-Time)  (cost=0.00..0.00 rows=0 width=0)
              ->  Distributed Subplan 3_1
                    ->  Limit  (cost=0.00..1.30 rows=10 width=847)
                          ->  Seq Scan on regions  (cost=0.00..13.00 rows=100 width=847)
              ->  Distributed Subplan 3_2
                    ->  Custom Scan (Citus Router)  (cost=0.00..0.00 rows=0 width=0)
                          Task Count: 1
                          Tasks Shown: All
                          ->  Task
                                Node: host=Node3 port=5432 dbname=sf1_0_3dgrid
                                ->  Limit  (cost=0.00..0.20 rows=10 width=44)
                                      ->  Seq Scan on periods_102012 periods  (cost=0.00..2.00 rows=100 width=44)
              Task Count: 32
              Tasks Shown: One of 32
              ->  Task
                    Node: host=Node2 port=5432 dbname=sf1_0_3dgrid
                    ->  Unique  (cost=20140805.49..20140981.89 rows=14112 width=40)
                          ->  Sort  (cost=20140805.49..20140840.77 rows=14112 width=40)
                                Sort Key: intermediate_result.regionid, intermediate_result_1.periodid, intermediate_result_1.period, c.licence
                                ->  Nested Loop  (cost=0.56..20139832.85 rows=14112 width=40)
                                      ->  Nested Loop  (cost=0.28..20135510.00 rows=14112 width=36)
                                            ->  Nested Loop  (cost=0.01..20010.01 rows=1000000 width=64)
                                                  ->  Function Scan on read_intermediate_result intermediate_result  (cost=0.00..10.00 rows=1000 width=36)
                                                  ->  Function Scan on read_intermediate_result intermediate_result_1  (cost=0.00..10.00 rows=1000 width=28)
                                            ->  Index Scan using trips_spgist_idx_102015 on trips_102015 t  (cost=0.28..20.11 rows=1 width=36)
                                                  Index Cond: (trip && stbox(intermediate_result.geom, intermediate_result_1.period))
                                                  Filter: ((trajectory(atperiod(trip, intermediate_result_1.period)) && intermediate_result.geom) AND _st_intersects(trajectory(atperiod(trip, intermediate_result_1.period)), intermediate_result.geom))
                                      ->  Index Scan using cars_pkey_102014 on cars_102014 c  (cost=0.28..0.31 rows=1 width=12)
                                            Index Cond: (carid = t.carid)
```

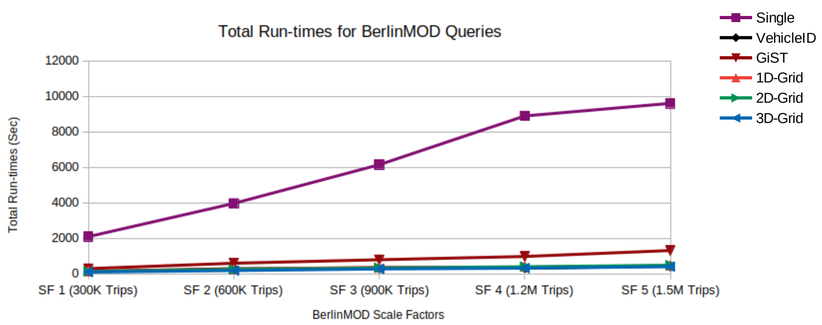
Distributed MobilityDB



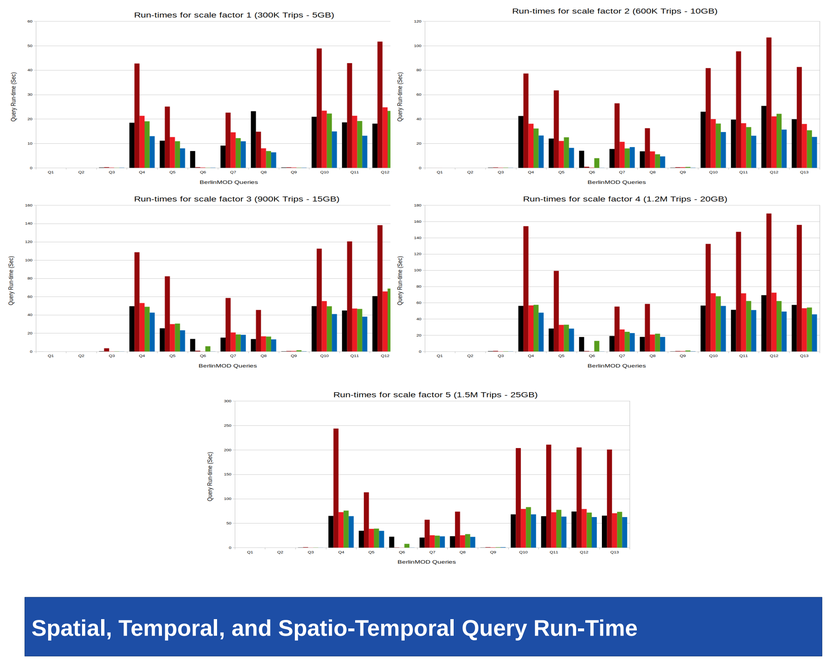
Experimental Dataset

Query	BerlinMOD Queries
Q1	What are the models of the vehicles with licence plate numbers from Licences?
Q2	How many vehicles exist that are passenger cars?
Q3	Where have the vehicles with licences from Licences been at each of the instants from Instants?
Q4	Which vehicles have passed the points from Points?
Q5	What are the licence plate numbers of the passenger cars that have reached the points from Points first of all passenger cars during the complete observation period?
Q6	What are the overall travelled distances of the vehicles with licence plate numbers from Licences during the periods from Periods?
Q7	What is the longest distance that was travelled by a vehicle during each of the periods from Periods?
Q8	Which vehicles passed a point from Points at one of the instants from Instants?
Q9	Which vehicles met at a point from Points at an instant from Instants?
Q10	Which vehicles travelled within one of the regions from Regions during the periods from Periods?
Q11	Which vehicles travelled within one of the regions from Regions at one of the instants from Instants?
Q12	Which vehicles passed a point from Points during a period from Periods?
Q13	Which point(s) from Points have been visited by a maximum number of different vehicles?

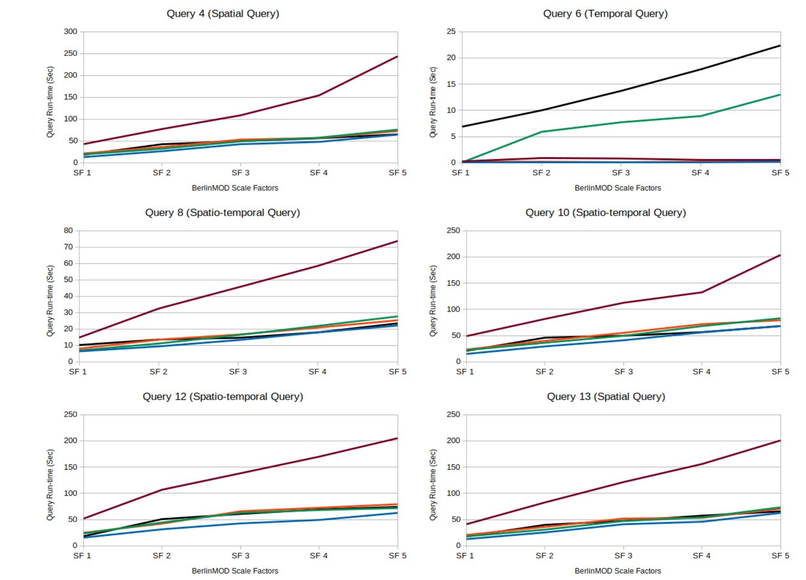
Partitioning Quality Measures



Queries Run-Time for Every Scale Factor



Spatial, Temporal, and Spatio-Temporal Query Run-Time



Tools

1.https://github.com/ULB-CoDE-WIT/MobilityDB
2.https://www.postgresql.org/
3.https://postgis.net/
4.https://www.citusdata.com/