X-Device Query Processing by Bitwise Distribution

Problem
- Spatial data points from tracked GPS devices
- 250M Points: (latitude, longitude, time, device)
- Reporting/Monitoring/Mining applications
- Many Spatial (Rectangular) Range Queries

Intermediates
- Approximation that can be used for, e.g., visualization
- May contain false positives, never false negatives
- Data Size reduction is high (factor 250K)

Decompose
- Vertical partitioning at bit-level
  - The GPU holds most significant bits
  - The CPU holds residuals
- Optimal partitioning is determined by analytical modeling

Compress
- Has to be lightweight and parallel decomposion friendly
  - Prefix Compression
  - Radix Clustering allows high compression ratio

Process
- GPU yields best effort approximations of the results.
  - Generates truncated values and false positives
  - Each thread of the GPU scans one cluster for one query
  - High parallelism provides high throughput.

Refine
- CPU copies partial results into main memory
- Positional joins reconstructs original values
- False positive are filtered out

Load Balance
- Overloaded GPU for single queries
- Almost perfect load balancing for larger query sets

Logos
- monetdb
- TELEIOS
- COMMIT