A Portable Relational Algebra Library for High-Performance Data-Intensive Query Processing
Ifrah Saeed, Haicheng Wu, and Sudhakar Yalamanchili
Georgia Institute of Technology

CHALLENGES

OUR SOLUTION

Candidate application domains

Diverse Architectures

Relational Computations Over Massive Unstructured Data

Applications and Software Stacks

Dataflow Diagram

Red Fox Platform Diagram

PERFORMANCE EVALUATION

Due to data transfer overhead, integrated GPUs and CPUs perform better than discrete GPUs for fine-grained primitives.

Due to higher parallelism and higher memory bandwidth, discrete GPUs outperform integrated devices for complex primitives.

When data fits in large CPU caches, CPUs perform better than other devices.

CONCLUSION

- Need for a portable library containing primitives that are required to execute data-intensive relational queries.
- When multiple devices are available in the system:
  - Schedule fine-grained primitives on the integrated GPU
  - Schedule complex primitives on the discrete GPU
  - Schedule primitives with cache sensitive footprints and large cache requirement on the multicore CPU

References
[6] Zhang, He, He, Lu, "Omnidb: Towards portable and efficient query processing on parallel CPU/GPU architectures." VLDB 2013