Red Fox: An Execution Environment for Relational Query Processing on GPUs

Haicheng Wu^, Gregory Diamos*, Tim Sheard #, Molham Aref ^, and Sudhakar Yalamanchili ! ^

^Georgia Institute of Technology, *NVIDIA Research, #Portland State University, ^LogicBlox Inc.

Kernel Weaver Module [4,5]

- Automatically perform Kernel Fusion Optimization
- Benefits:
  - Reduce Data Movements
  - Reduce Memory Footprints
  - Enlarge Compiler Optimization Scope

If fusing below operators together on Tesla C2070,

Then get additional 2.89x speedup on average.

Red Fox TPC-H Benchmark Performance

(GTX GeForce Titan)[7]

- Execution time = PCIe + GPU Computation
- Problem size is restricted by GPU memory capacity
- No data movement optimizations
- Unoptimized query plan
- JOIN and SORT are from ModernGPU Library: Set, Unique, Aggregation are from Thrust library. The others are designed by us.
- Each Operator maps to several CUDA Kernels

Compare with LogicBlox 4.0 on 2 8-core CPUs (SF = 1)

7x Faster with 1/3 Price

Most time spent in SORT

Current Work: Scaling to out-of-core data sets in Accelerator Clouds

References


MICRO 2012

Example Datalog Query

1. number(n) :- int32(n).
2. number(0).
3. if other facts elided for brevity
4. next(n,m) :- int32(n), int32(m).
5. next(0,1).
6. if other facts elided for brevity
7. even(n) :- int32(n).
8. even(n) :- int32(n).
9. even(0).
10. even(n) :- int32(n), odd(m).
11. even(n) :- int32(n).
12. odd(n) :- int32(n).
13. odd(n) :- int32(n), even(m).

Example Query Plan (CFG)

BB1:
- pre_odd := odd
- pre_even := even
- odd := Join next even (x1)

BB2:
- m1 := MapFilter next
- s1 := Sort Join number m1
- even := MapJoin s1 odd (x1)

BB3:
- if pre_odd = odd?

BB4:
- if pre_even = even?

BB5:
- HALT

Recursive Definition

Translatable to Loops

Red Fox: An Execution Environment for Relational Query Processing on GPUs

Haicheng Wu^, Gregory Diamos*, Tim Sheard #, Molham Aref ^, and Sudhakar Yalamanchili ! ^

^Georgia Institute of Technology, *NVIDIA Research, #Portland State University, ^LogicBlox Inc.

Kernel Weaver Module [4,5]

- Automatically perform Kernel Fusion Optimization
- Benefits:
  - Reduce Data Movements
  - Reduce Memory Footprints
  - Enlarge Compiler Optimization Scope

If fusing below operators together on Tesla C2070,

Then get additional 2.89x speedup on average.

Red Fox TPC-H Benchmark Performance

(GTX GeForce Titan)[7]

- Execution time = PCIe + GPU Computation
- Problem size is restricted by GPU memory capacity
- No data movement optimizations
- Unoptimized query plan
- JOIN and SORT are from ModernGPU Library: Set, Unique, Aggregation are from Thrust library. The others are designed by us.
- Each Operator maps to several CUDA Kernels

Compare with LogicBlox 4.0 on 2 8-core CPUs (SF = 1)

7x Faster with 1/3 Price

Most time spent in SORT

Current Work: Scaling to out-of-core data sets in Accelerator Clouds

References


MICRO 2012

Example Datalog Query

1. number(n) :- int32(n).
2. number(0).
3. if other facts elided for brevity
4. next(n,m) :- int32(n), int32(m).
5. next(0,1).
6. if other facts elided for brevity
7. even(n) :- int32(n).
8. even(n) :- int32(n).
9. even(0).
10. even(n) :- int32(n), odd(m).
11. even(n) :- int32(n).
12. odd(n) :- int32(n).
13. odd(n) :- int32(n), even(m).

Example Query Plan (CFG)

BB1:
- pre_odd := odd
- pre_even := even
- odd := Join next even (x1)

BB2:
- m1 := MapFilter next
- s1 := Sort Join number m1
- even := MapJoin s1 odd (x1)

BB3:
- if pre_odd = odd?

BB4:
- if pre_even = even?

BB5:
- HALT

Recursive Definition

Translatable to Loops

Red Fox: An Execution Environment for Relational Query Processing on GPUs

Haicheng Wu^, Gregory Diamos*, Tim Sheard #, Molham Aref ^, and Sudhakar Yalamanchili ! ^

^Georgia Institute of Technology, *NVIDIA Research, #Portland State University, ^LogicBlox Inc.

Kernel Weaver Module [4,5]

- Automatically perform Kernel Fusion Optimization
- Benefits:
  - Reduce Data Movements
  - Reduce Memory Footprints
  - Enlarge Compiler Optimization Scope

If fusing below operators together on Tesla C2070,

Then get additional 2.89x speedup on average.