

Polystore Optimization via Program Expression Graphs

Dylan Hutchison* Bill Howe Dan Suciu
Zachary Tatlock University of Washington

Polystore Optimization: to synthesize an optimal multi-system program from an input script's specification.



Matrix Multiply Example – a Saturated PEG¹

1. Parse the input script into an initial PEG
2. Add equivalences to algebras (Raco², Lara³, Array) and available systems (Myria, CombBLAS)
3. Select the lowest cost execution plan, which may involve multiple systems

Vision for Systems-free Programming: Specify computation declaratively. Optimizers will synthesize the best execution plan on available systems w.r.t. equivalences and a cost model.

Input Script, written in MyriaL

```
A = scan(tableA); B = scan(tableB);  
C = select A.r, B.c, sum(A.v * B.v)  
  from A, B where A.c = B.r;  
D = select r, c, 2*v from C;  
export(D, fileD, 'csv');
```

Challenges

- Cost models
- Writing rules
- Access Control
- Provenance

