TIMO SCHNEIDER <TIMOS@INF.ETHZ.CH>

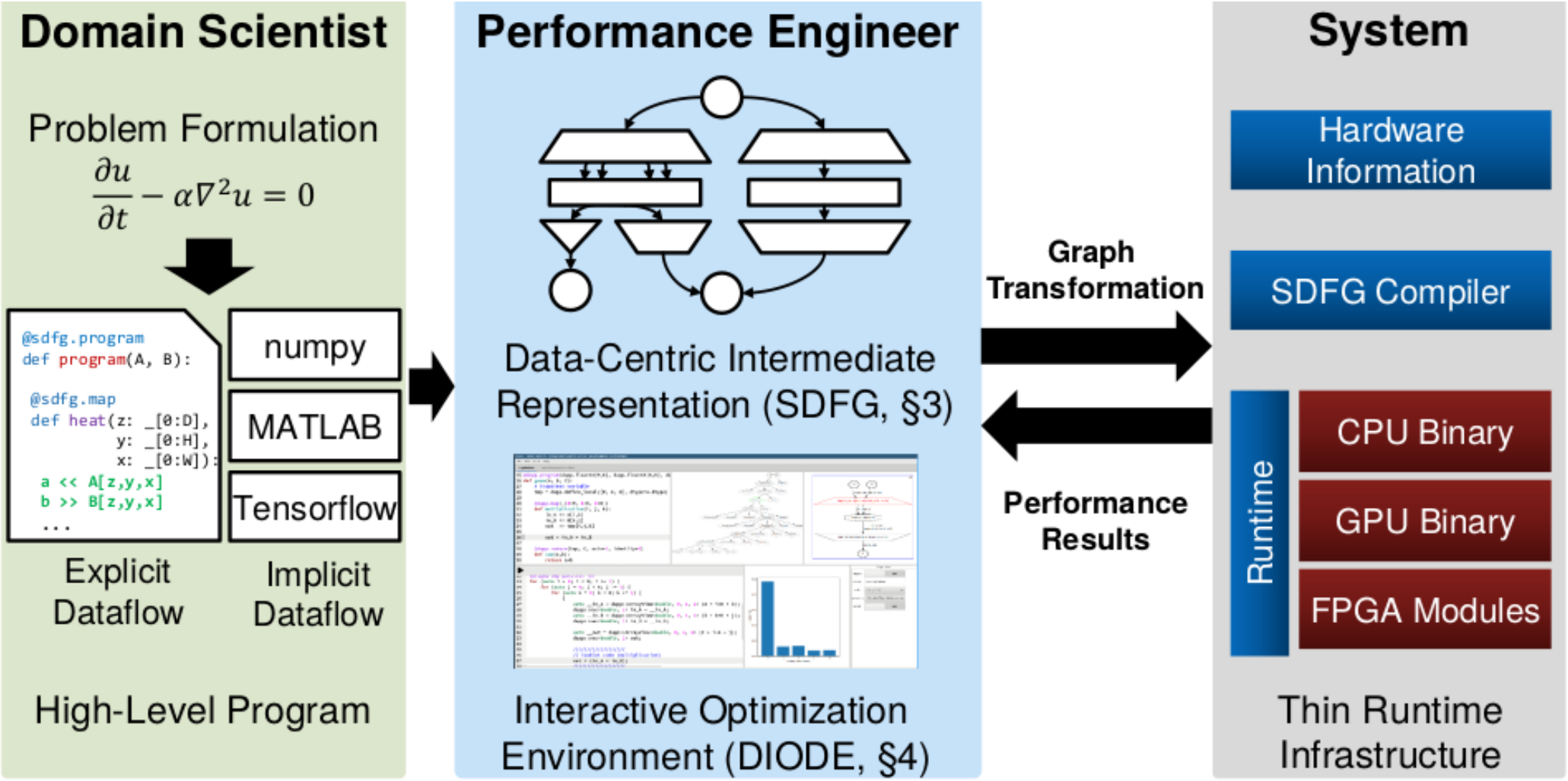
Master Thesis: Distributed Memory Support for Dataflow Programming

Thesis Topic

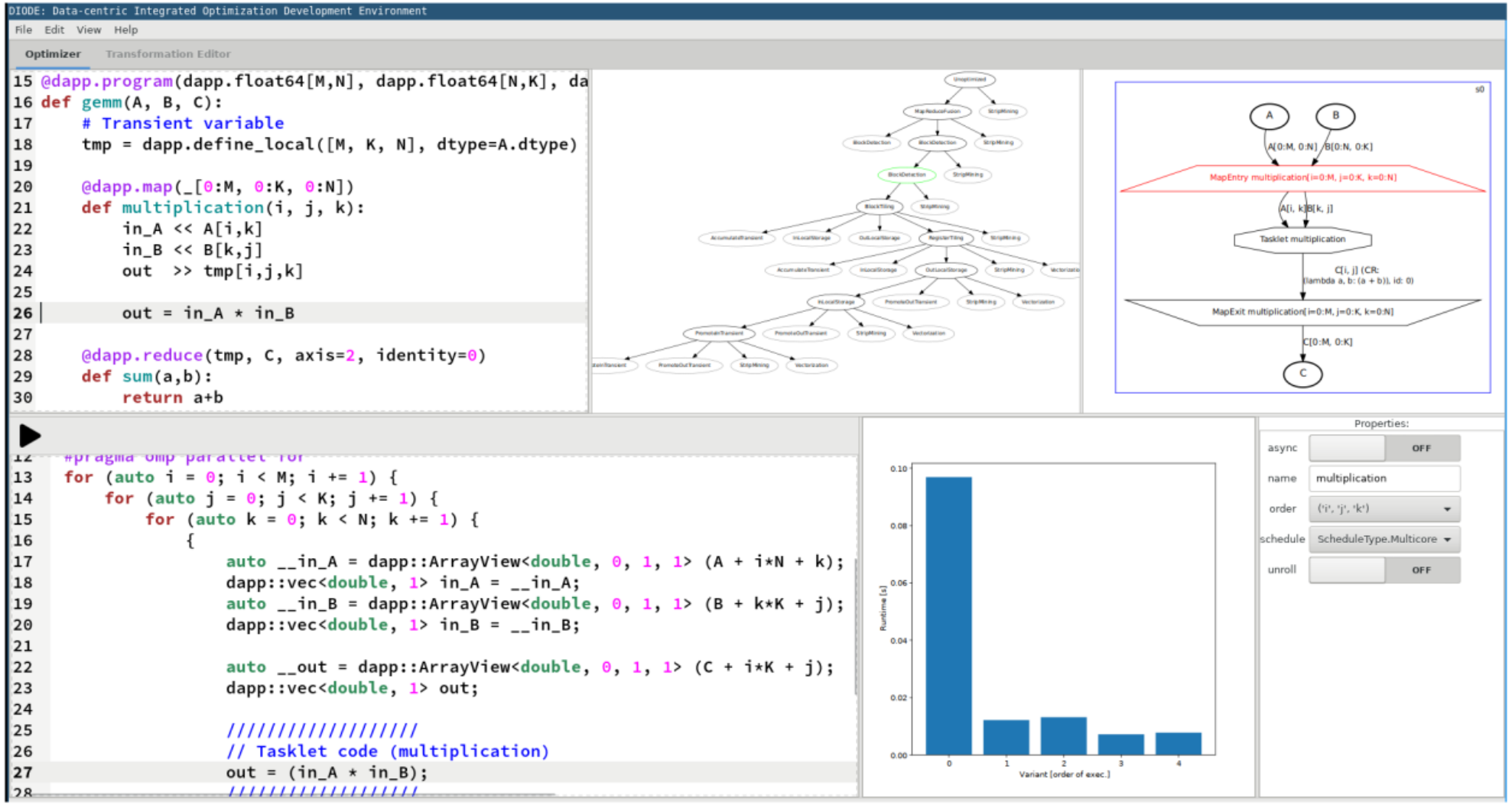
- We developed a graph-based representation for parallel programs
- We currently lower to multicore CPUs, GPUs, and FPGAs
- Your task would be: integrate support for distributed memory machines
- Idea is to use MPI RMA

- You already know about parallel programming, MPI, etc.
- You will work with a team of PhD students and post-docs
- You will work on a large code-base

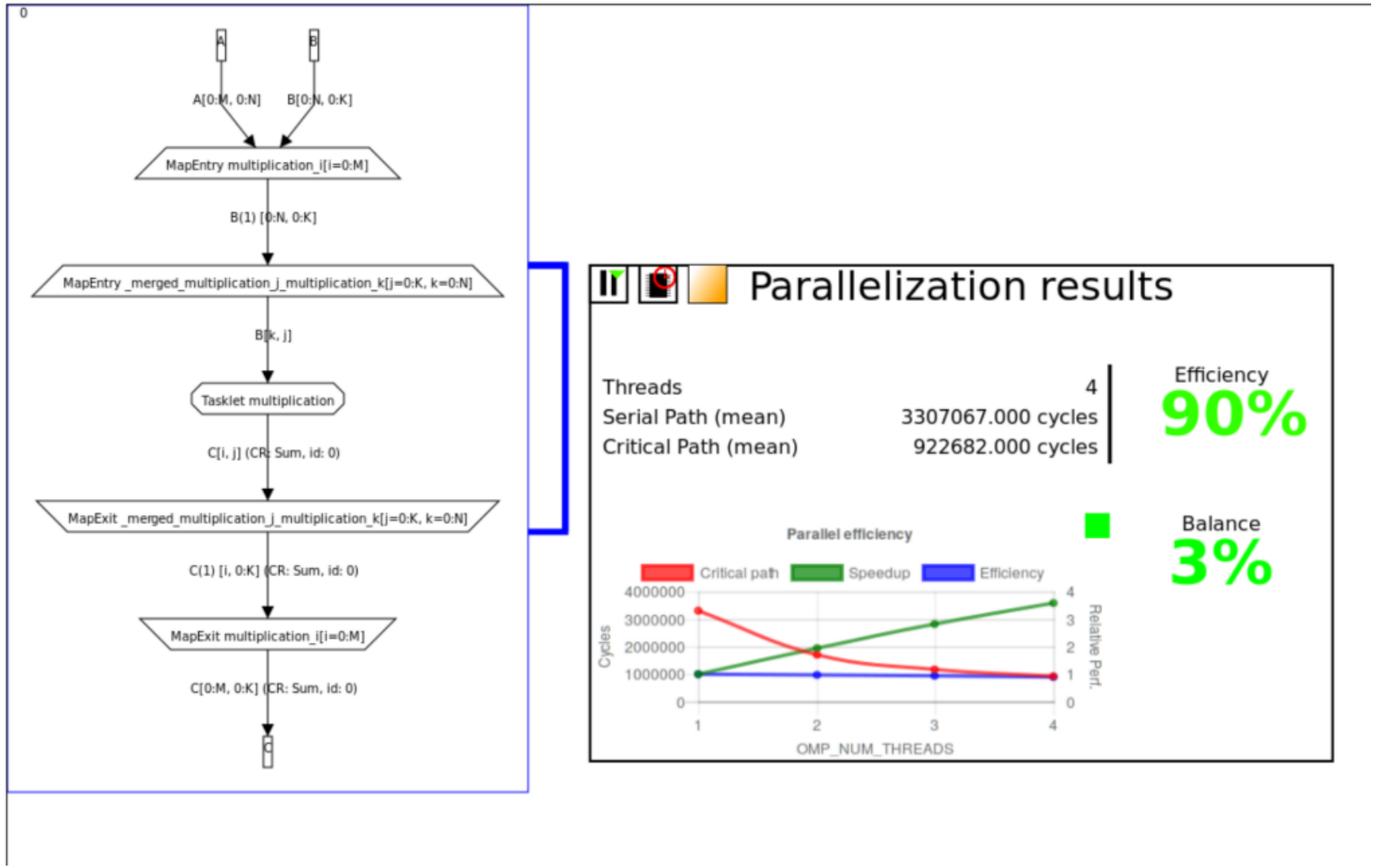
The Big Picture



DIODE: The Future of Performance Tuning?

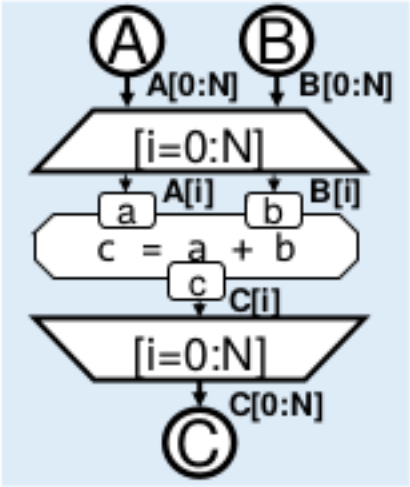


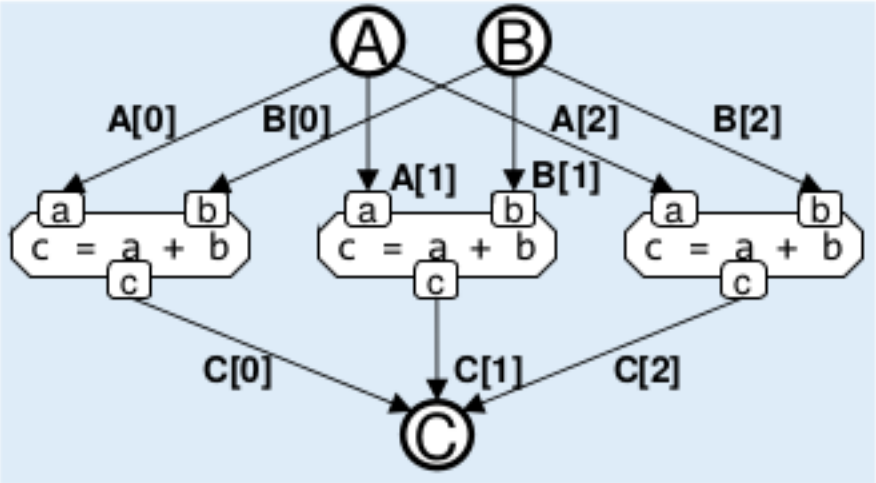
DIODE: The Future of Performance Tuning?



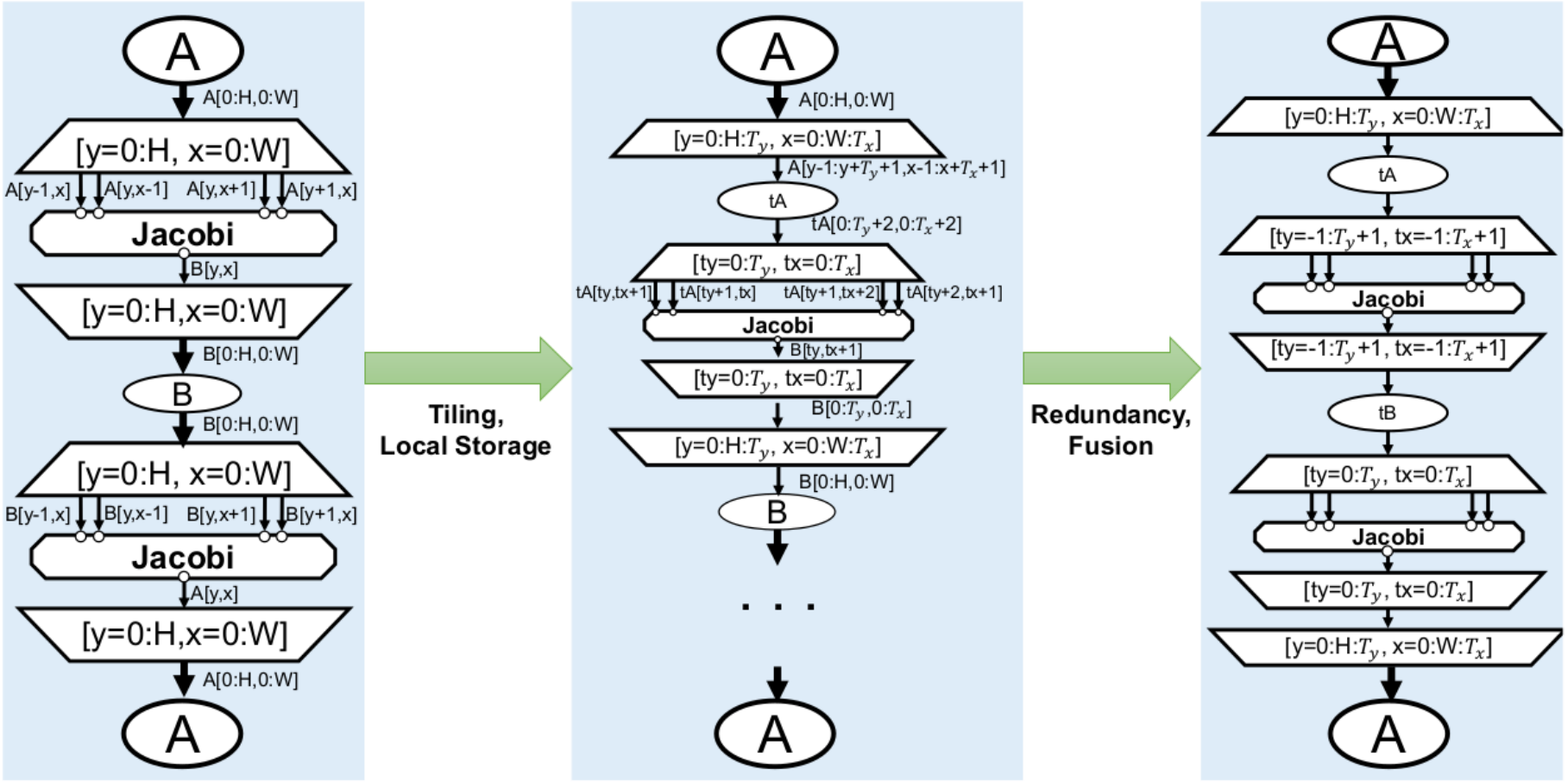
Stateful Dataflow Multigraphs (SDFGs)

Primitive	Description
Turing-Complete Construction	
s0 — iter < N → s1	**State**: State machine element.
Data, Transient Data	**Data**: N-dimensional array container.
Tasklet	**Tasklet**: Fine-grained computational block.
A[i,j] → B(1) [0:M,k] →	**Memlet**: Data movement descriptor.
Dataflow and Concurrency	
Stream	**Stream**: Streaming data container.
[i=0:M, j=0:N], [i=0:M, j=0:N]	**Map**: Parametric graph abstraction for parallelism.
C [i,j] (CR: Sum) →	**Conflict Resolution**: Defines behavior during conflicting writes.
Nesting and Subgraph Aliases	
Invoke	**Invoke**: Call a nested SDFG.
sum, id: 0	**Reduce**: Reduction of one or more memlet axes.
[p=0:P], s>0, [p=0:P], s>0	**Consume**: Dynamic mapping of computations on stream elements.





Optimizing SDFGs via Graph Transformations

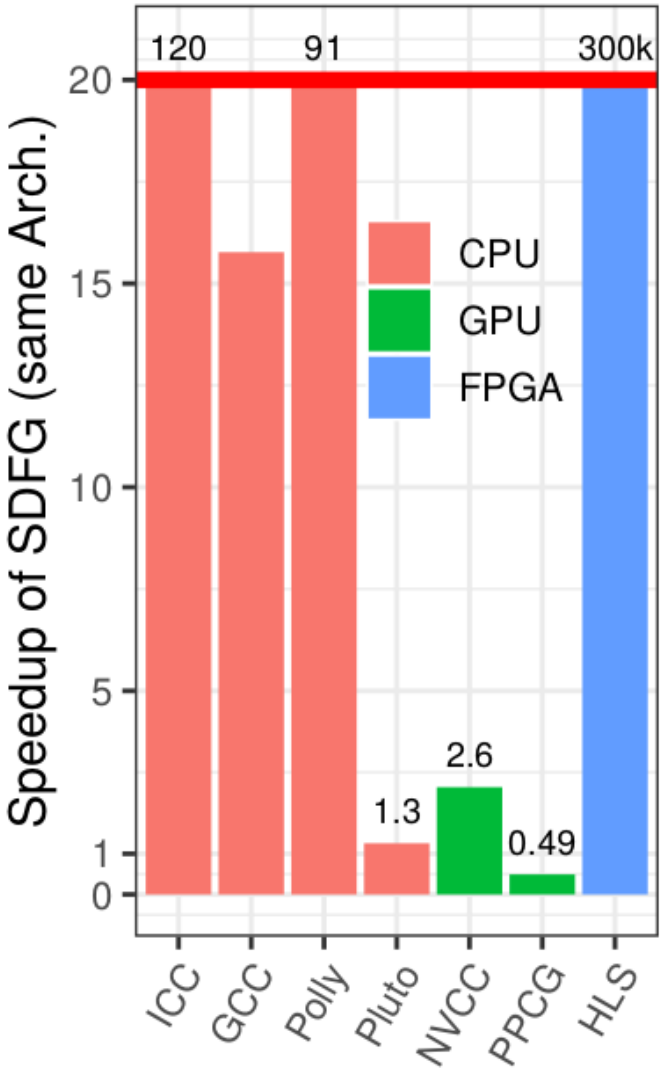


Memory: $2WH$
Performance: **6 GFLOP/s**

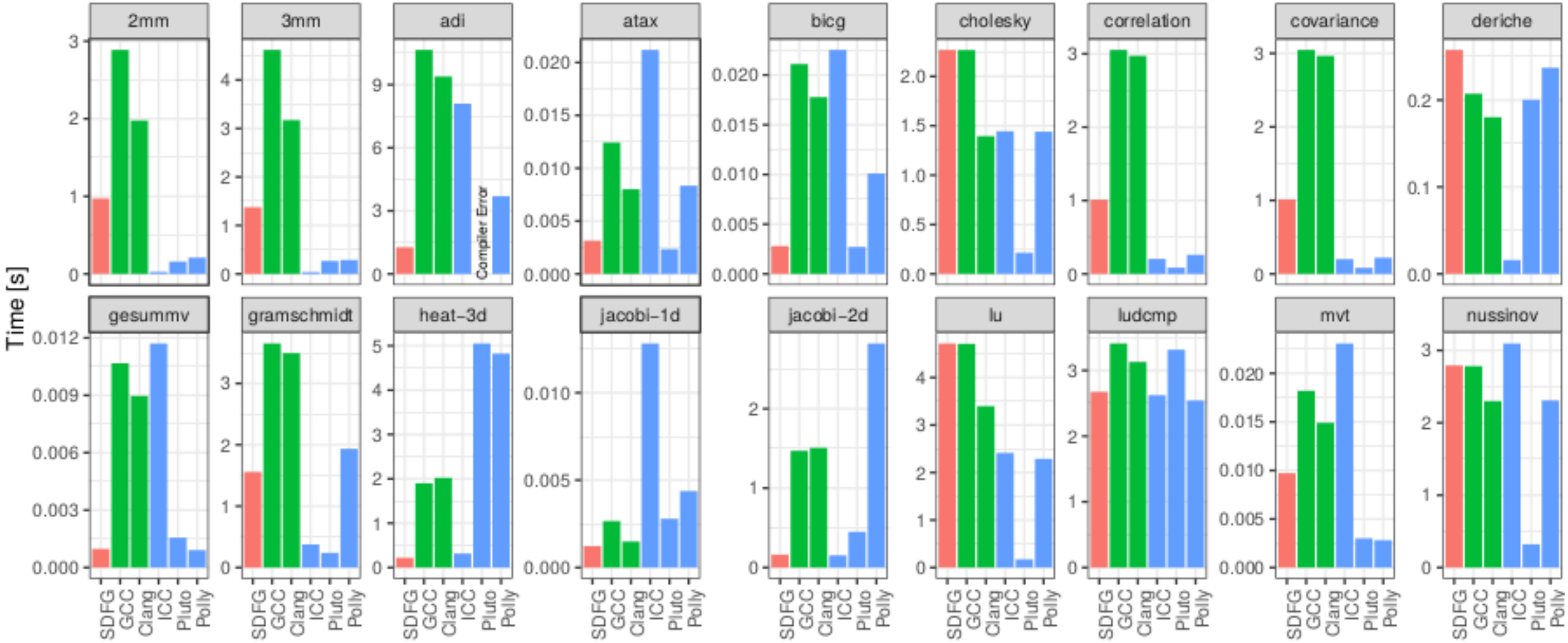
Memory: $WH + 2T_xT_y$
Perf.: **22 GFLOP/s**

Video of DIODE in action (tuning MM): https://vimeo.com/301317247

So does this really work?



So does this really work?



What you need to do

- Data-nodes need to be annotated with a storage location
- When reading from non-local memory, generate MPI_Get()
- Add synchronization in the right place
- Collective “detection”
- Look at lots of example SDFGs and the generated code

- Details are flexible of course, if interested just contact us!

timos@inf.ethz.ch