# AMNESIAC: Amnesic Automatic Computer

## Trading Computation for Communication for Energy Efficiency

**Ismail Akturk** and Ulya R. Karpuzcu

University of Minnesota, Twin Cities

{aktur002, ukarpuzc}@umn.edu

ALTAI

UNIVERSITY OF MINNESOTA

# Motivation

| Technology Node | 40nm | 10nm |
|---|---|---|
| Energy of on chip communication normalized to computation | 1.55x | 5.77x |

Keckler, S. W., Dally, W. J., Khailany, B., Garland, M., and Glasco, D. “GPUs and the Future of Parallel Computing”, IEEE Micro 31, 5 (2011).



# Motivation

| Technology Node | 40nm | 10nm |
|---|---|---|
| Energy of on chip communication normalized to computation | 1.55x | 5.77x |

**(Re)computing becomes cheaper than memorizing and retrieving data**

Keckler, S. W., Dally, W. J., Khailany, B., Garland, M., and Glasco, D. “GPUs and the Future of Parallel Computing”, IEEE Micro 31, 5 (2011).



# Amnesic* Automatic Computer

- Idea: Replace energy-hungry load with sequence of instructions

*amnesia: noun a partial or total loss of memory (origin late 18th cent.: from Greek amnēsia ‘forgetfulness’)

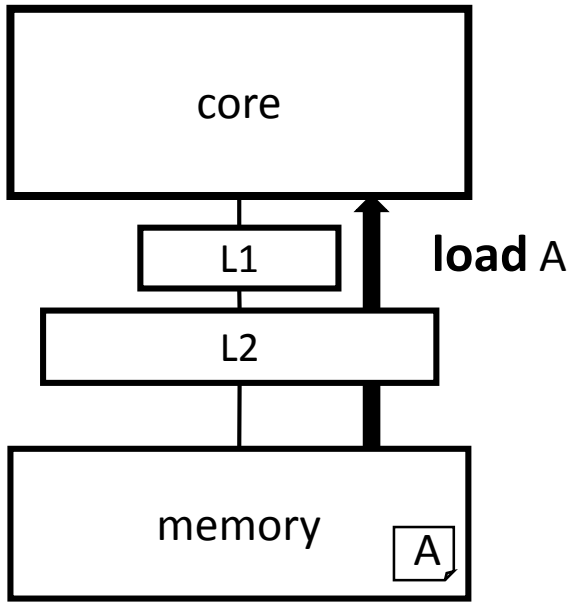
# Amnesic* Automatic Computer

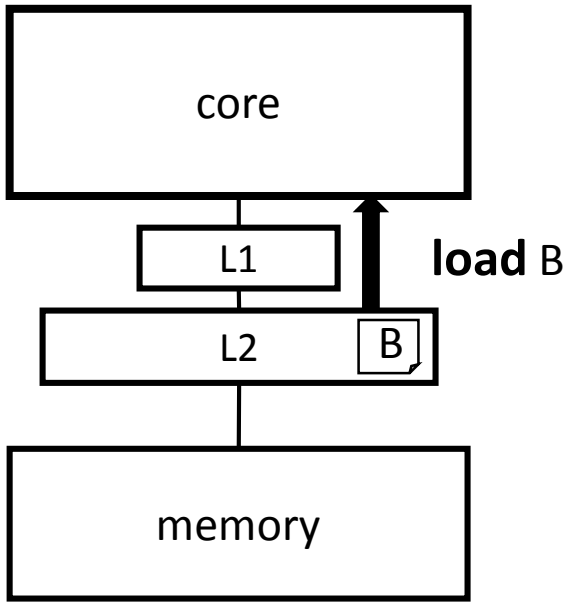
- Idea: Replace energy-hungry load with sequence of instructions
- Scope
  - How to determine which data values to recompute
  - How to identify instructions to recompute data values
  - How to protect architectural state during recomputation
  - How to orchestrate recomputation at runtime

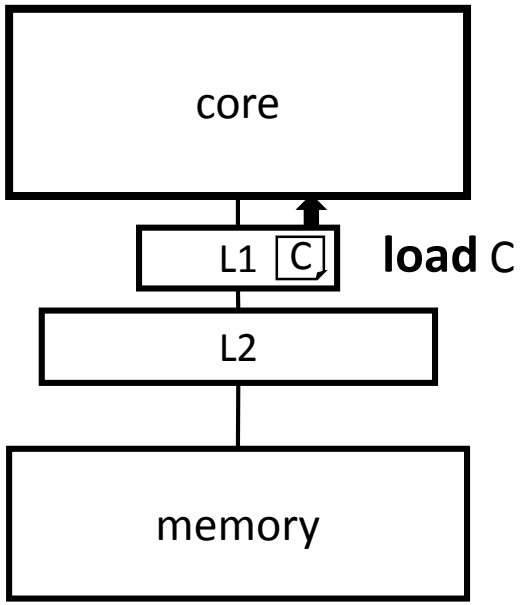
*amnesia: noun a partial or total loss of memory (origin late 18th cent.: from Greek amnēsia ‘forgetfulness’)



# How to determine which data values to recompute



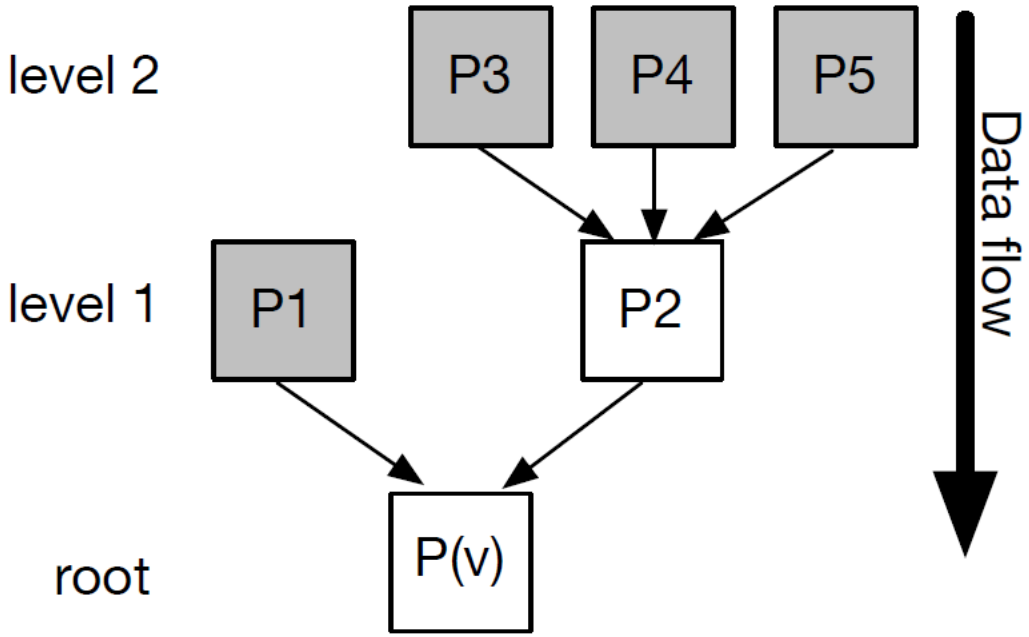




$E_{ld,A}$ ? $E_{rc,A}$

$E_{ld,B}$ ? $E_{rc,B}$

$E_{ld,C}$ ? $E_{rc,C}$

# Recomputation Slice (RSlice)



# Overview

- Amnesic Compiler
  - Identifies independent RSlices
  - Annotates RSlices: RCMP, REC, RTN



# Overview

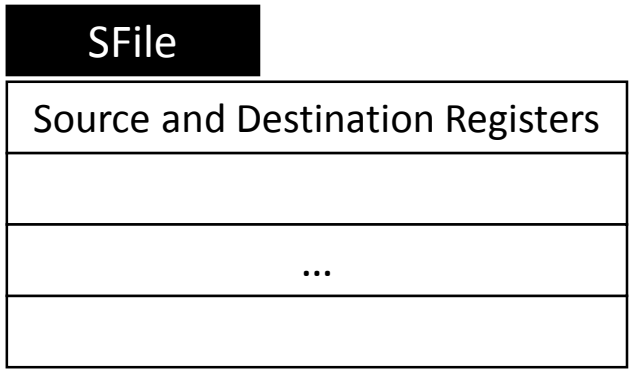
- Amnesic Compiler
  - Identifies independent RSlices
  - Annotates RSlices: RCMP, REC, RTN
- Amnesic Scheduler
  - Decides when to recompute
  - Can use various policies
    - Compiler
    - First-Level Cache Miss - FLC
    - Last-Level Cache Miss - LLC



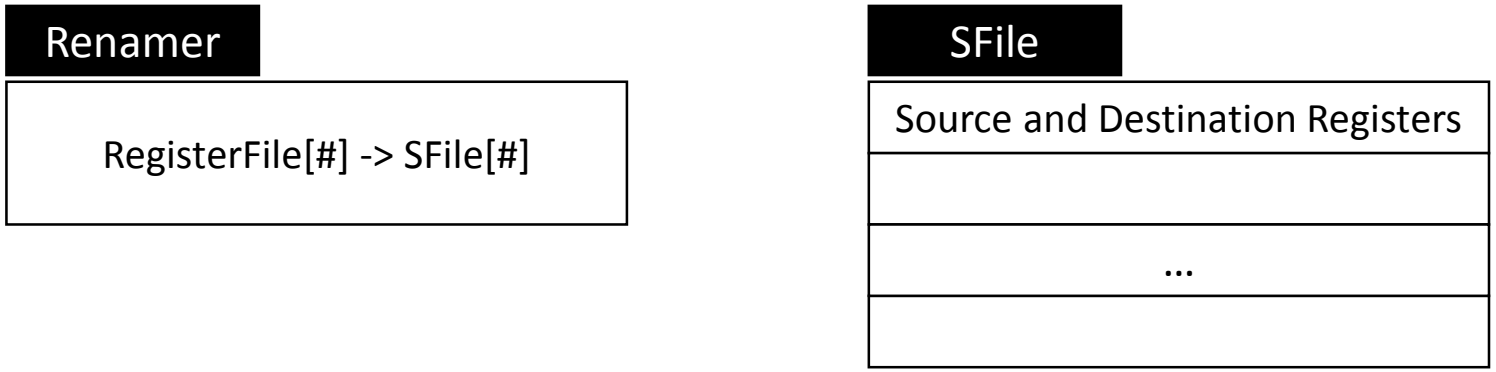
# Overview

- Amnesic Compiler
  - Identifies independent RSlices
  - Annotates RSlices: RCMP, REC, RTN
- Amnesic Scheduler
  - Decides when to recompute
  - Can use various policies
    - Compiler
    - First-Level Cache Miss - FLC
    - Last-Level Cache Miss - LLC
- Amnesic Microarchitecture
  - Prevents corruption of architectural state
  - Makes RSlice inputs available at the time of recomputation

# Amnesic Microarchitecture



# Amnesic Microarchitecture



# Amnesic Microarchitecture

**Renamer**

| RegisterFile[#] -> SFile[#] |
|---|

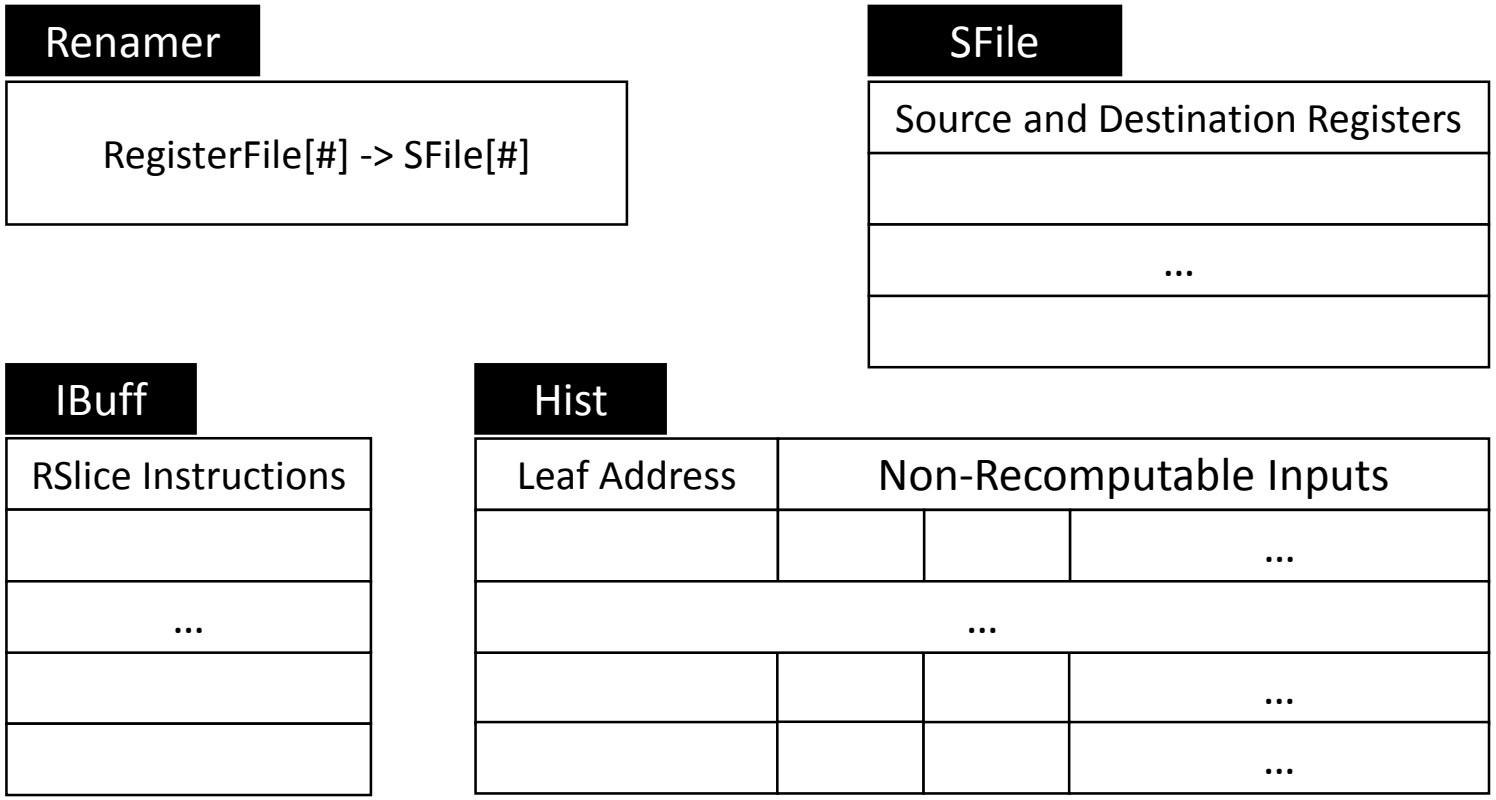
**SFile**

| Source and Destination Registers |
|---|
| |
| … |
| |

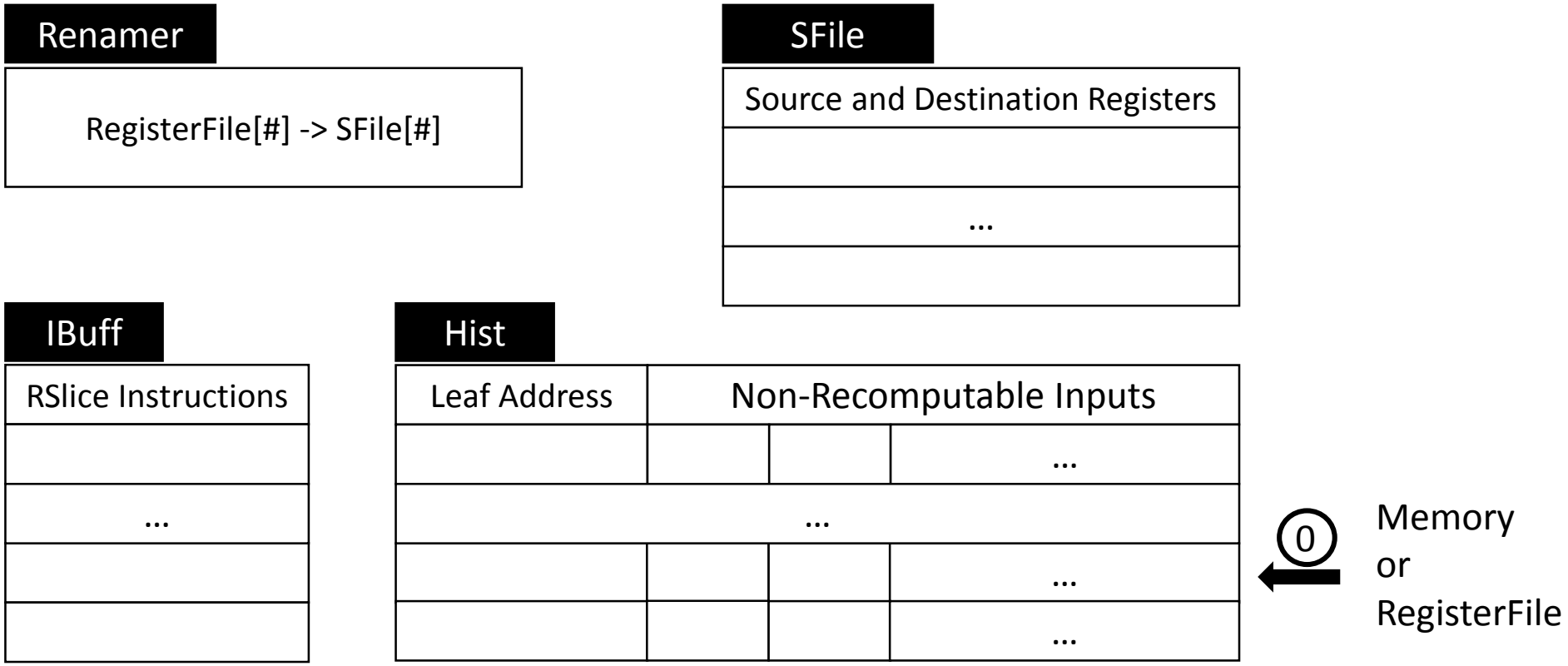
**Hist**

| Leaf Address | Non-Recomputable Inputs | | |
|---|---|---|---|
| | | | … |
| … | | | |
| | | | … |
| | | | … |

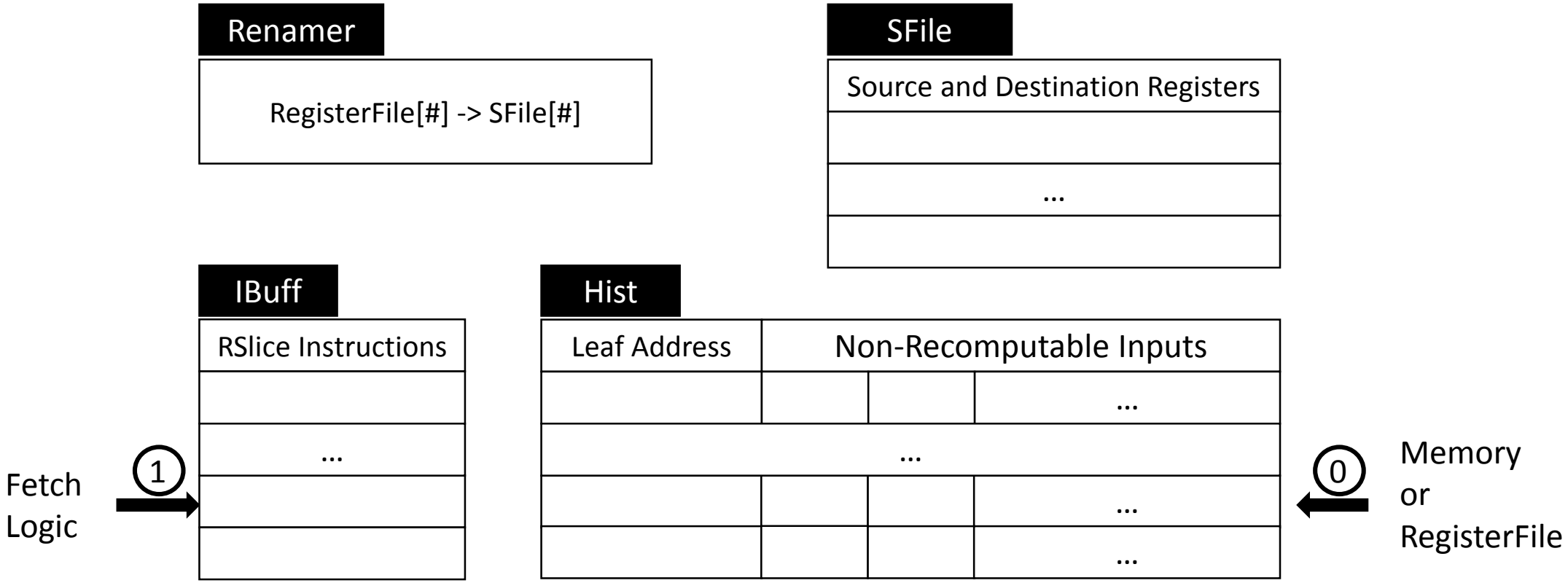
# Amnesic Microarchitecture



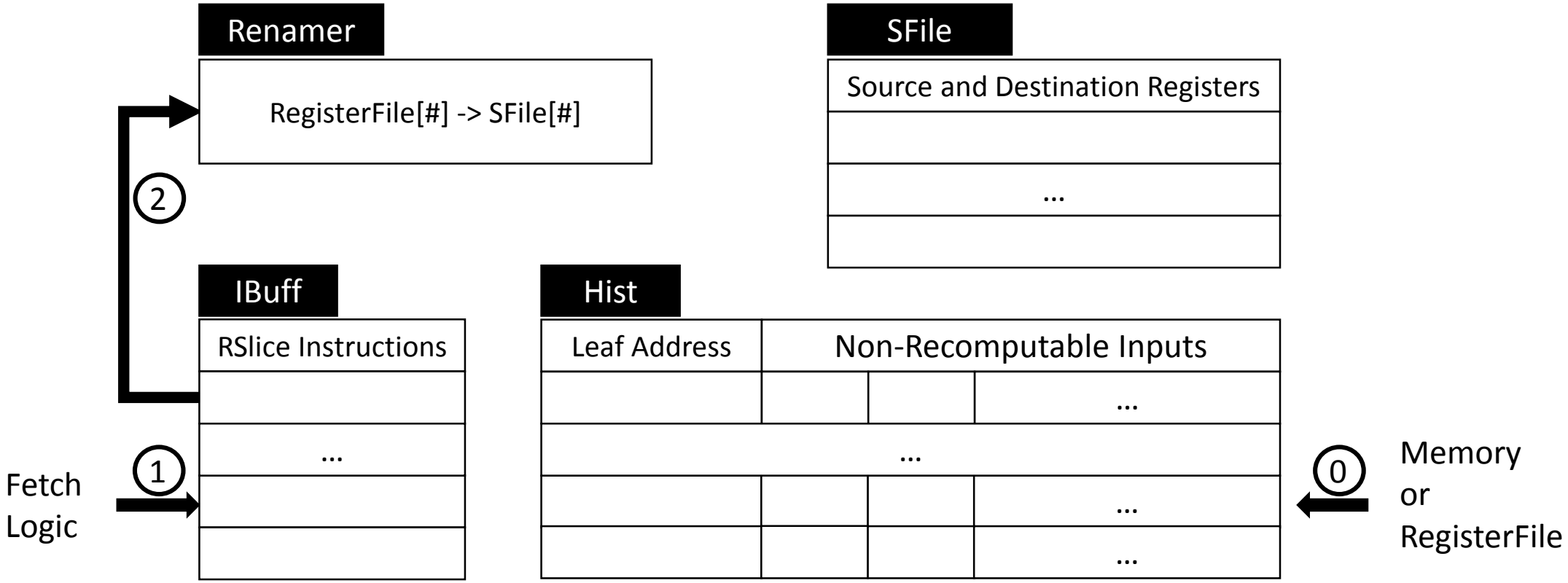
# Putting It All Together: Default Execution



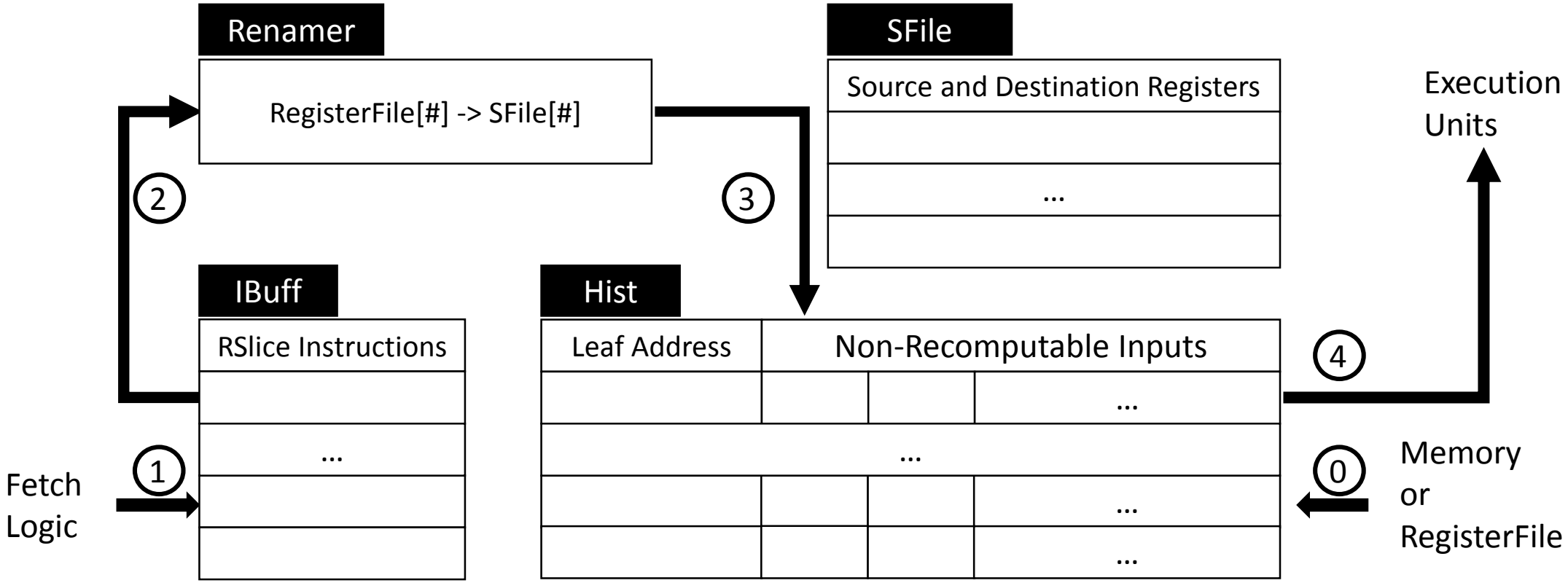
# Putting It All Together: Default Execution



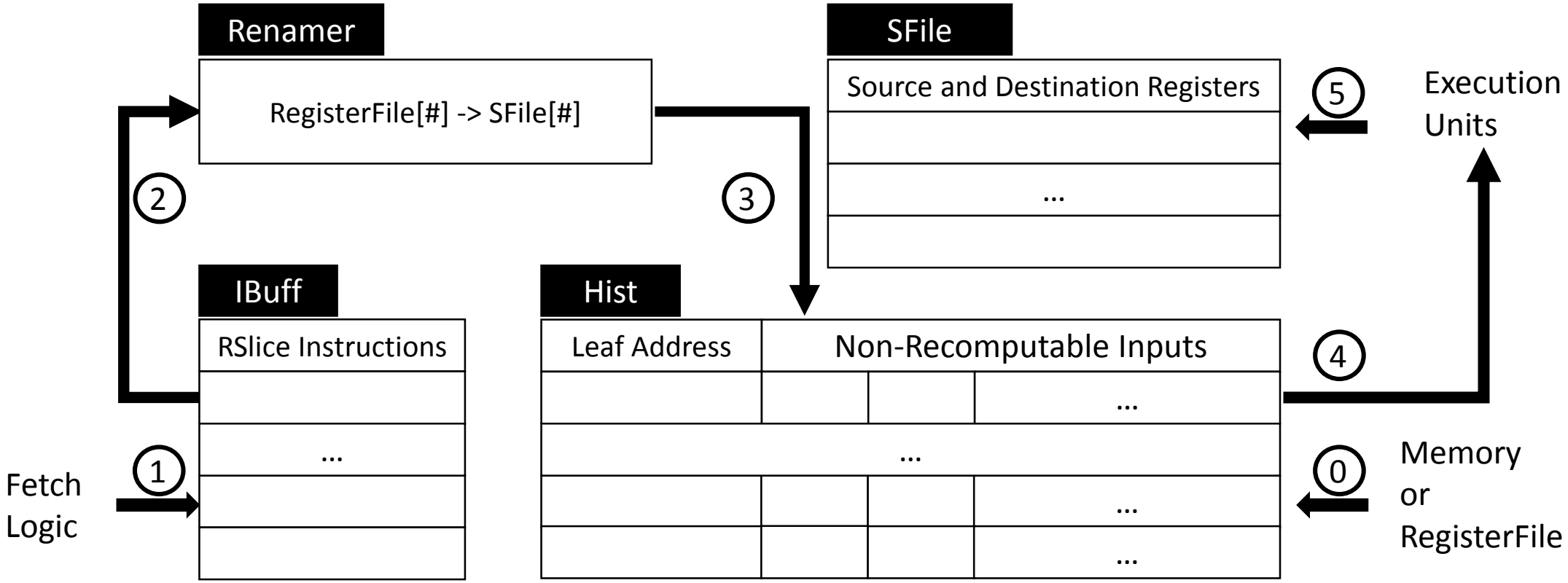
# Putting It All Together: During Recomputation



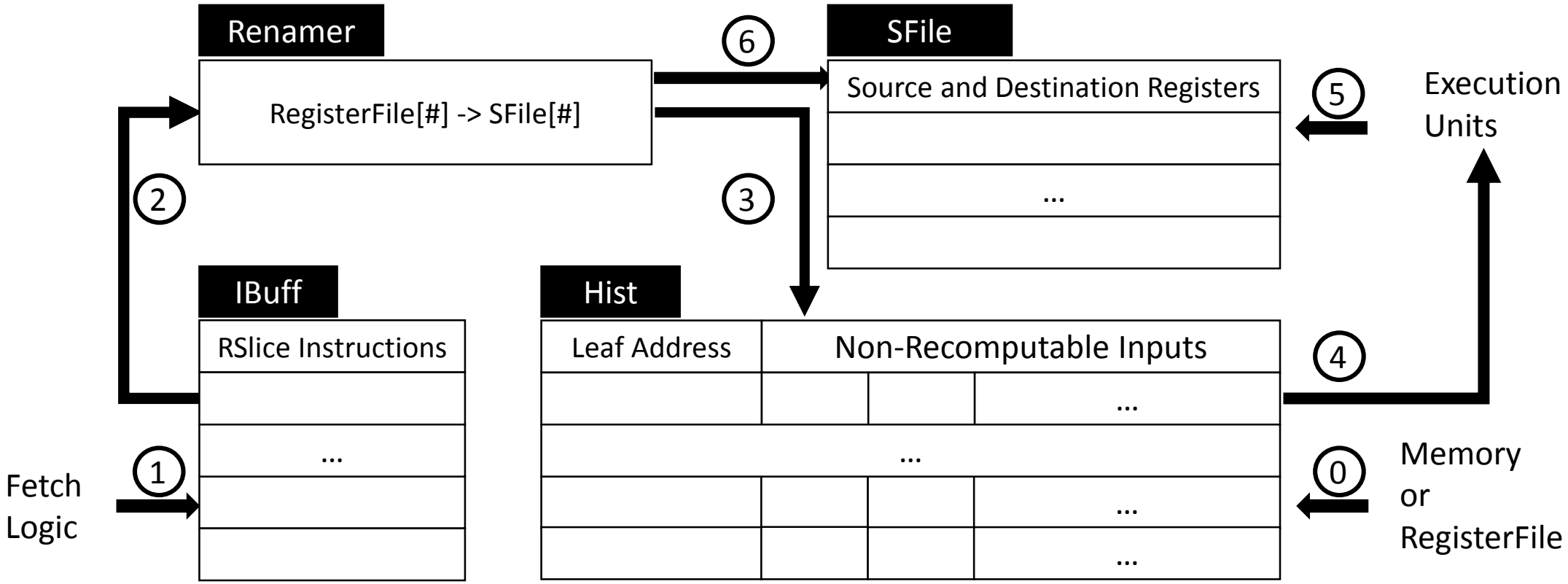
# Putting It All Together: During Recomputation



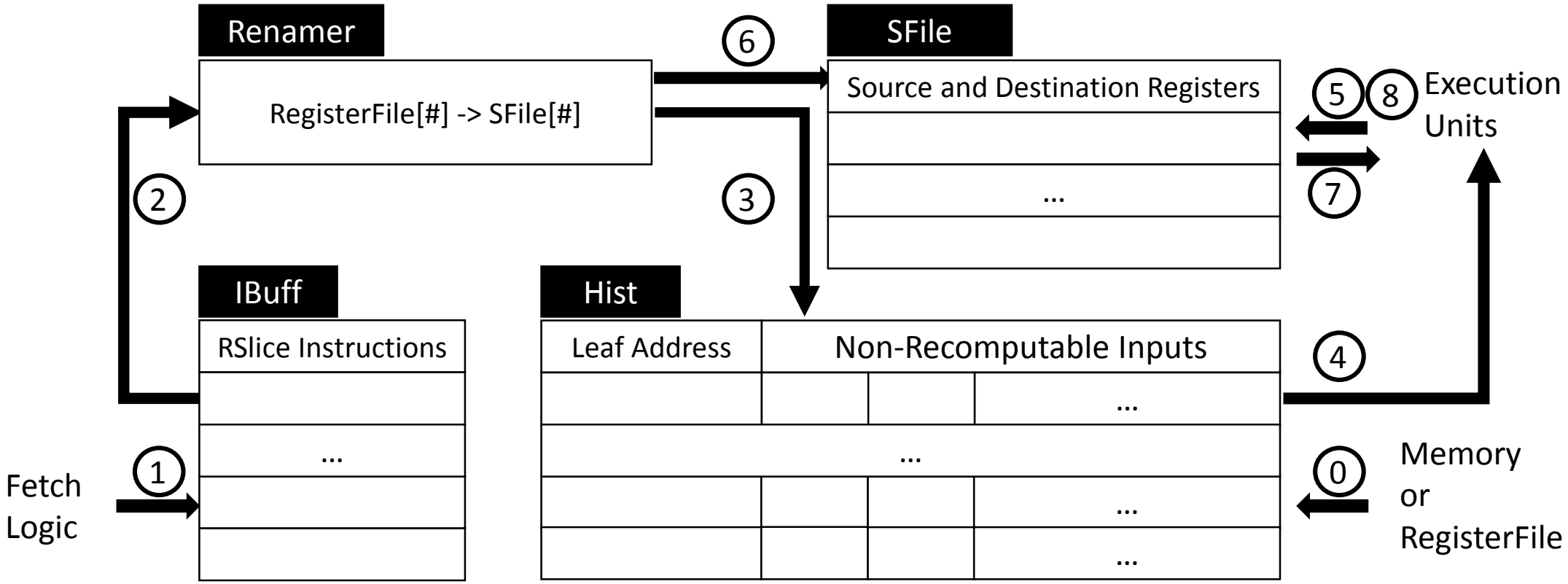
# Putting It All Together: During Recomputation



# Putting It All Together: During Recomputation



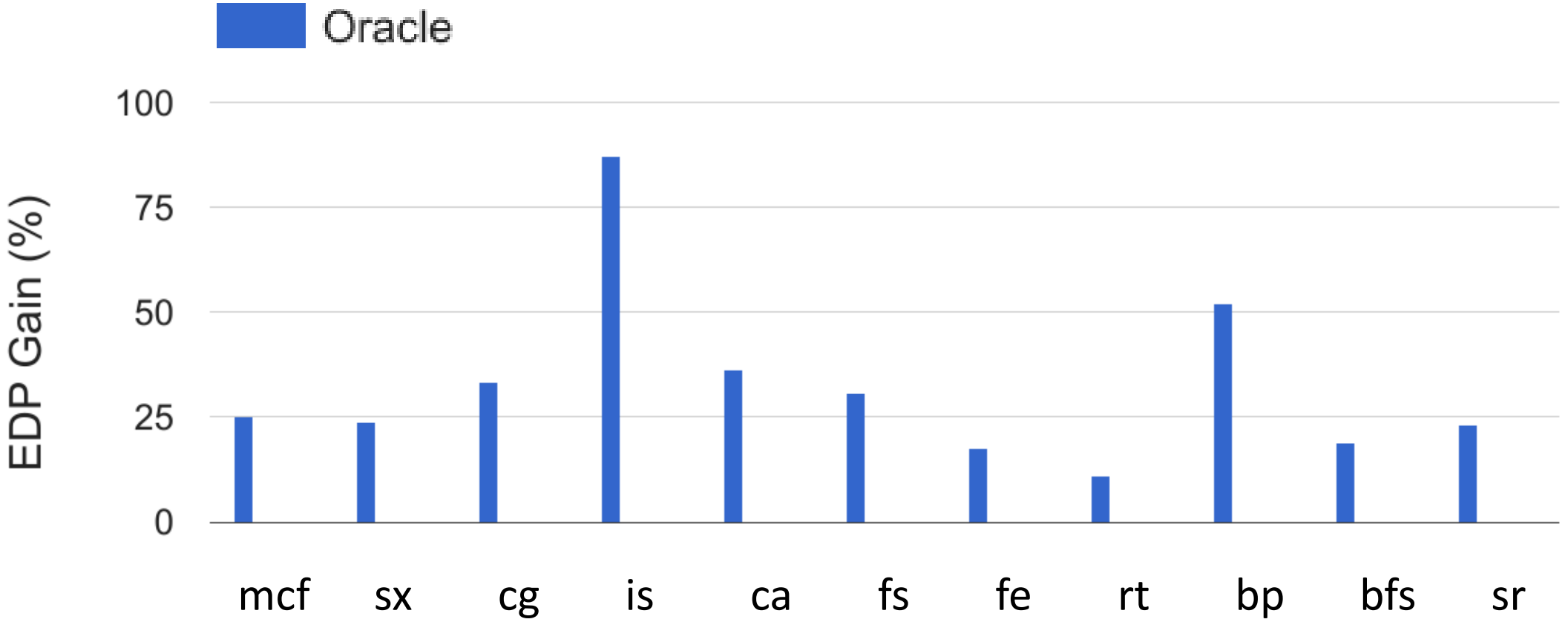
# Putting It All Together: During Recomputation



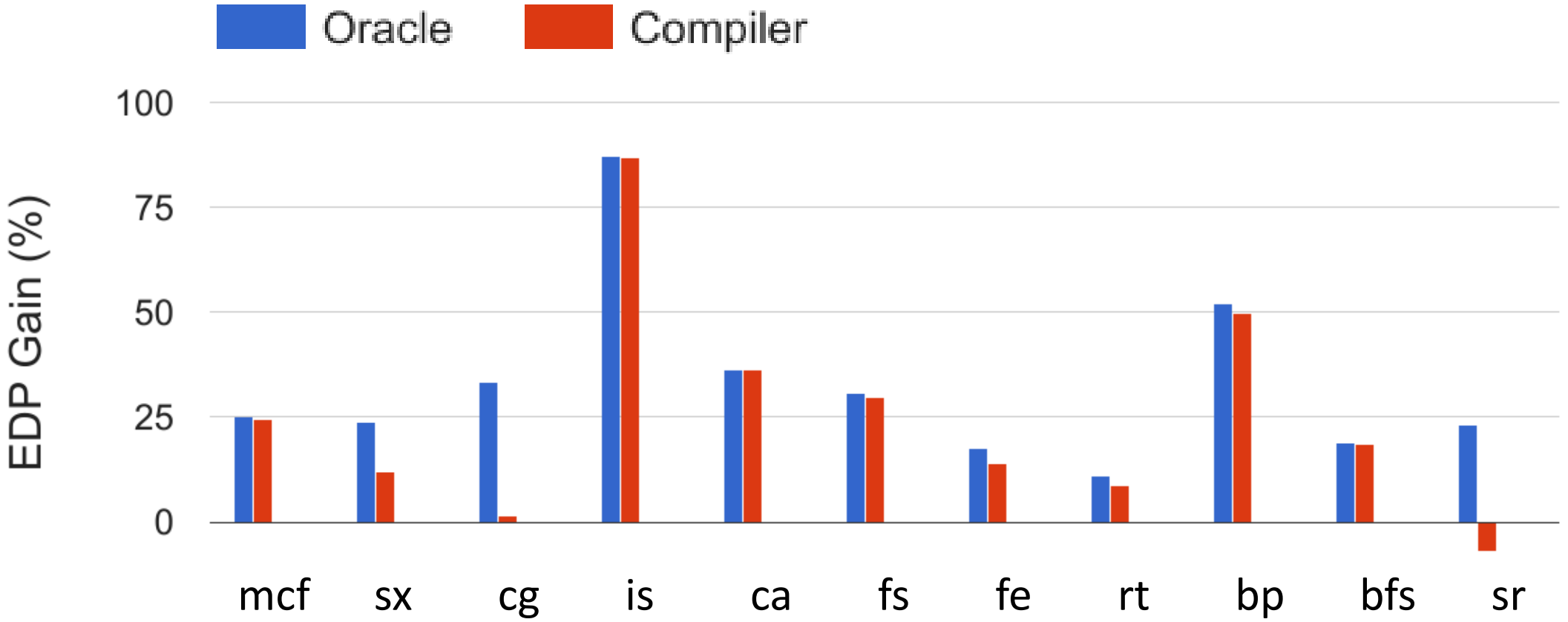
# Evaluation Setup

- Benchmarks: SPEC2006, NAS, Parsec, Rodinia
  - 33 benchmarks evaluated, 11 of them show > 10% EDP gain
- Amnesic compiler pass mimicked by binary generator Pintool
- Microarchitecture and scheduler implemented in Snipersim

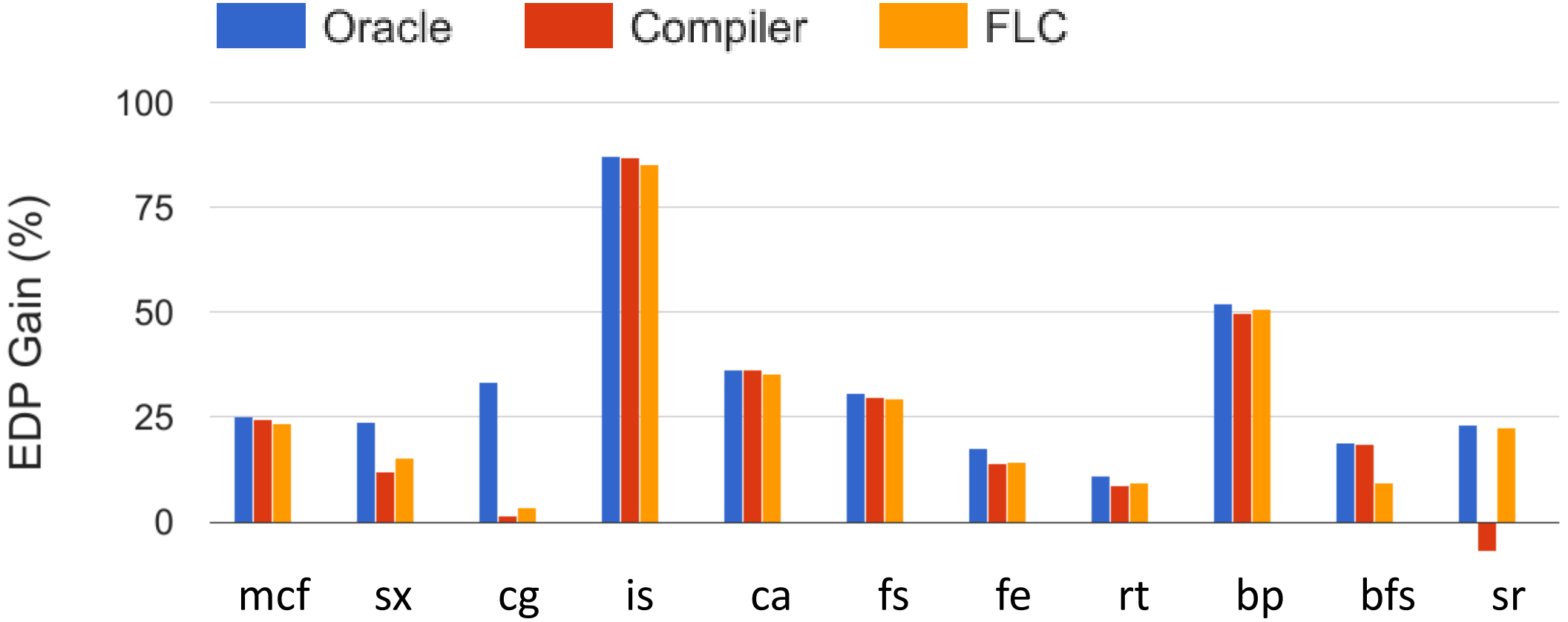
# Energy-Delay Product



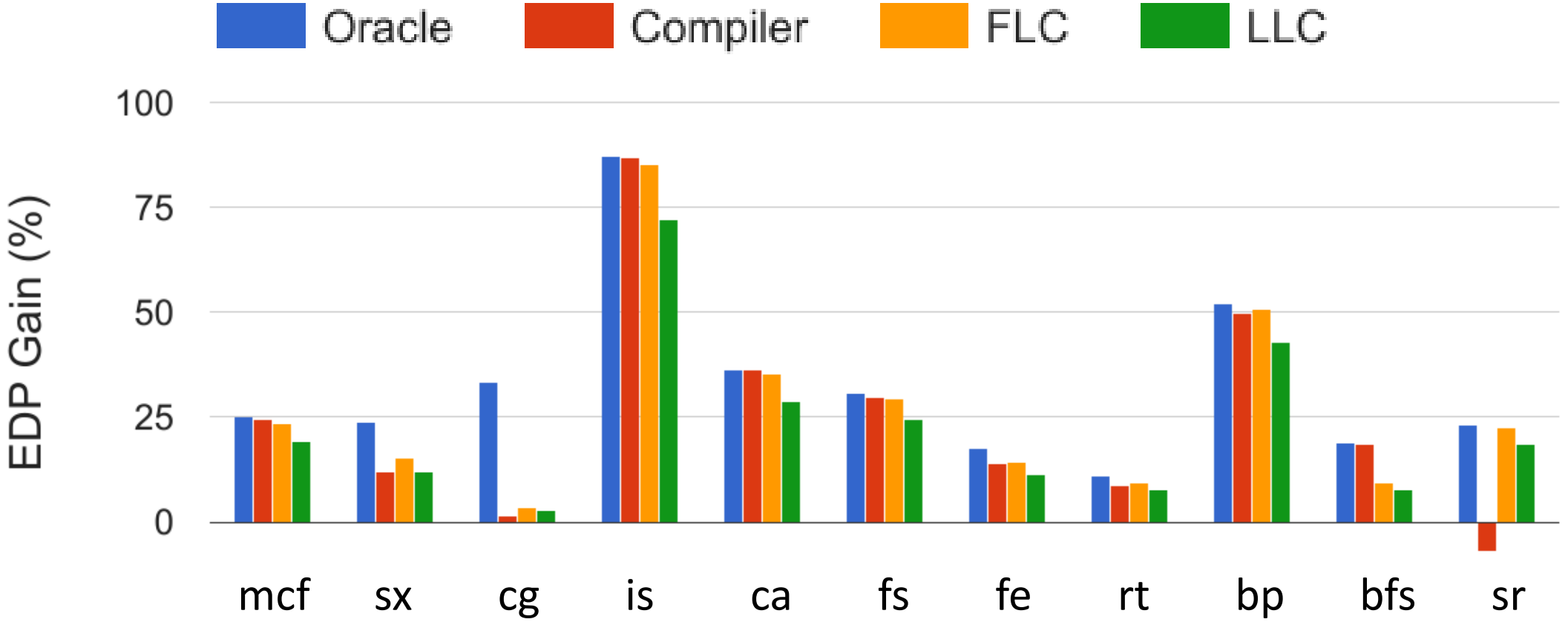
# Energy-Delay Product



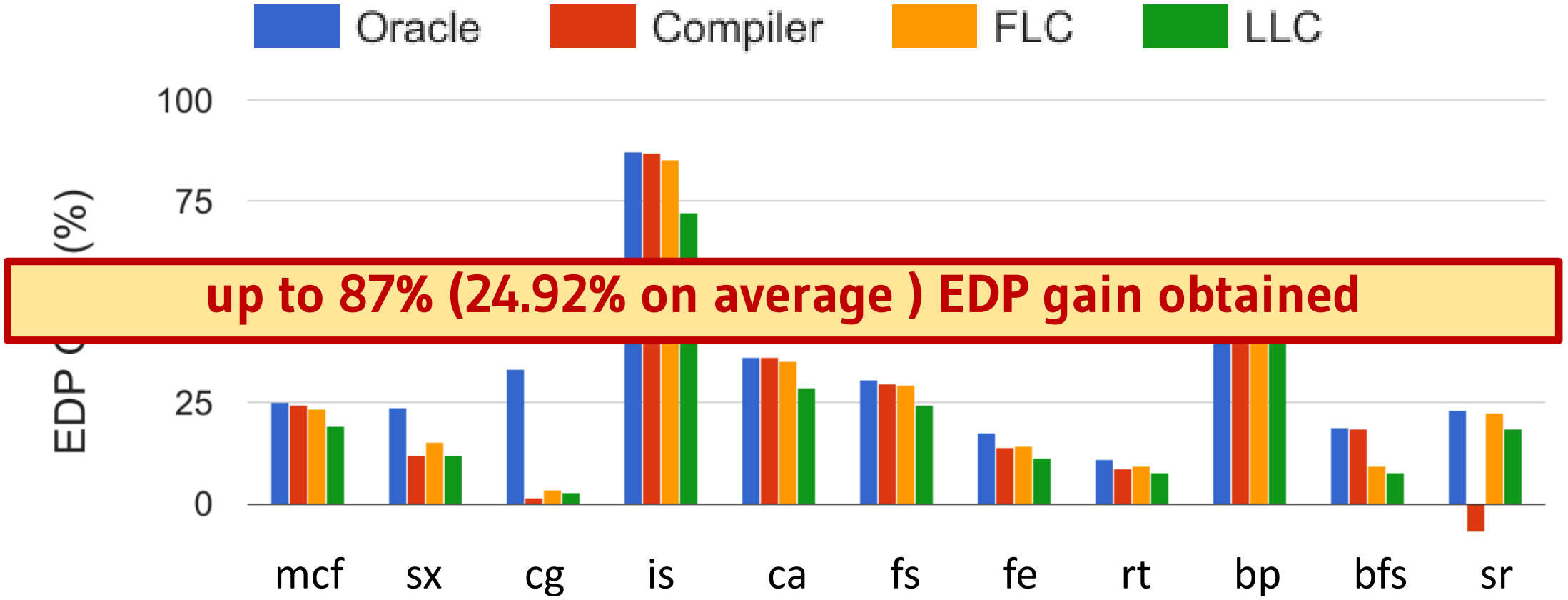
# Energy-Delay Product



# Energy-Delay Product



# Energy-Delay Product




**up to 87% (24.92% on average ) EDP gain obtained**

# Instruction Mix

| Benchmark | % increase in instruction count | % decrease in load count |
|---|---|---|
| mcf | 4.47 | 6.19 |
| sx | 4.55 | 6.68 |
| cg | 3.97 | 2.11 |
| is | 17.97 | 49.99 |
| ca | 7.38 | 7.95 |
| fs | 1.83 | 3.08 |
| fe | 3.55 | 1.75 |
| rt | 1.97 | 6.08 |
| bp | 31.89 | 55.55 |
| bfs | 1.20 | 60.93 |
| sr | 20.02 | 23.33 |



# Instruction Mix

| Benchmark | % increase in instruction count | % decrease in load count |
|---|---|---|
| mcf | 4.47 | 6.19 |
| is | 17.97 | 49.99 |
| ca | 7.38 | 7.95 |
| fs | 1.83 | 3.08 |
| fe | 3.55 | 1.75 |
| rt | 1.97 | 6.08 |
| bp | 31.89 | 55.55 |
| bfs | 1.20 | 60.93 |
| sr | 20.02 | 23.33 |

**49.99% reduction in loads, at the expense of 17.97% more instructions**



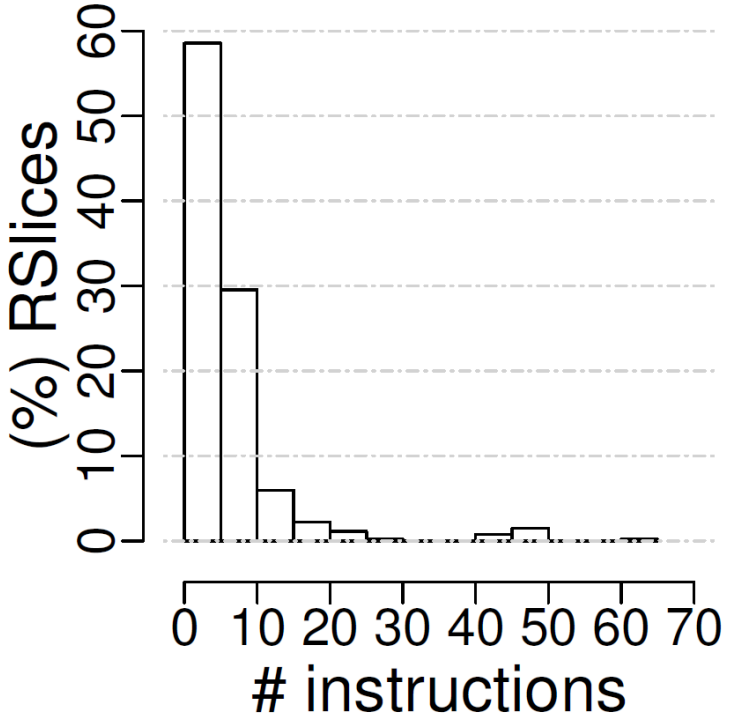
# Instruction Mix

| Benchmark | % increase in instruction count | % decrease in load count |
|---|---|---|
| mcf | 4.47 | 6.19 |
| is | 17.97 | 49.99 |
| fe | 3.55 | 1.75 |
| rt | 1.97 | 6.08 |
| bp | 31.89 | 55.55 |
| bfs | 1.20 | 60.93 |
| sr | 20.02 | 23.33 |

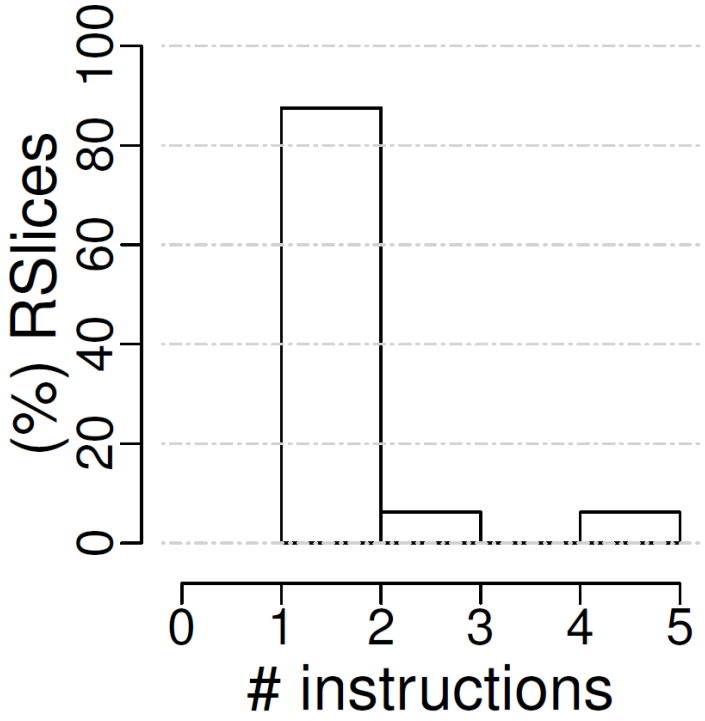
**49.99% reduction in loads, at the expense of 17.97% more instructions**

**translates into 74.68% reduction in (load) energy**

# RSlice Length

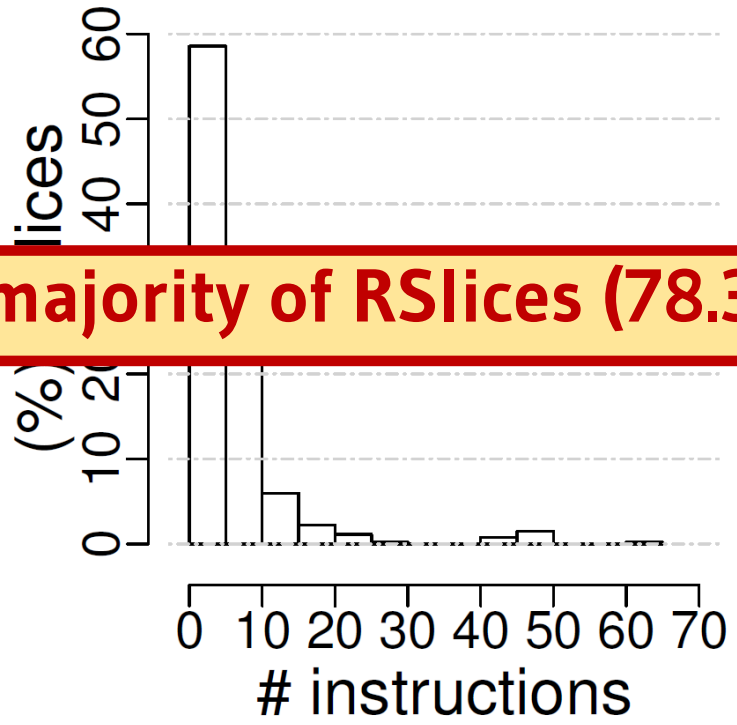



sx

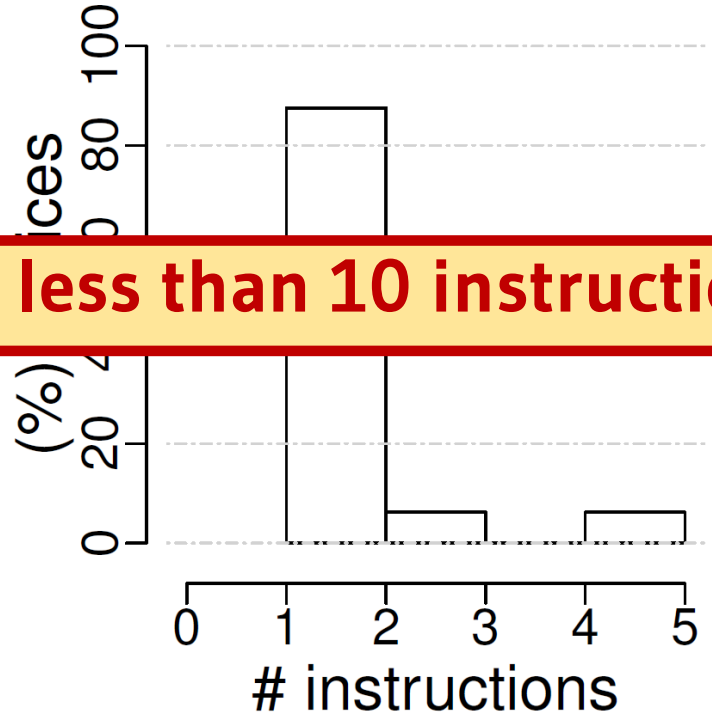



bfs

# RSlice Length




sx




bfs

**majority of RSlices (78.32%) have less than 10 instructions**

# Summary

- Amnesic execution can minimize power and performance overhead of
  - Data retrieval
  - Data communication
- Amnesic execution makes the workload more compute-intensive
  - Better use of classic architectures
- Up to 87% improvement in energy-delay product (avg. 24%)



# Questions and Comments

Please send your questions and comments to

aktur002@umn.edu

# AMNESIAC: Amnesic Automatic Computer

## Trading Computation for Communication for Energy Efficiency

**Ismail Akturk** and Ulya R. Karpuzcu

University of Minnesota, Twin Cities

{aktur002, ukarpuzc}@umn.edu

ALTAI

UNIVERSITY OF MINNESOTA