Partitioning Register Files to Reduce Access Time

Kyle Bryson, John Kim,
Supratik Majumder, Julie Rosser

April 22, 2003

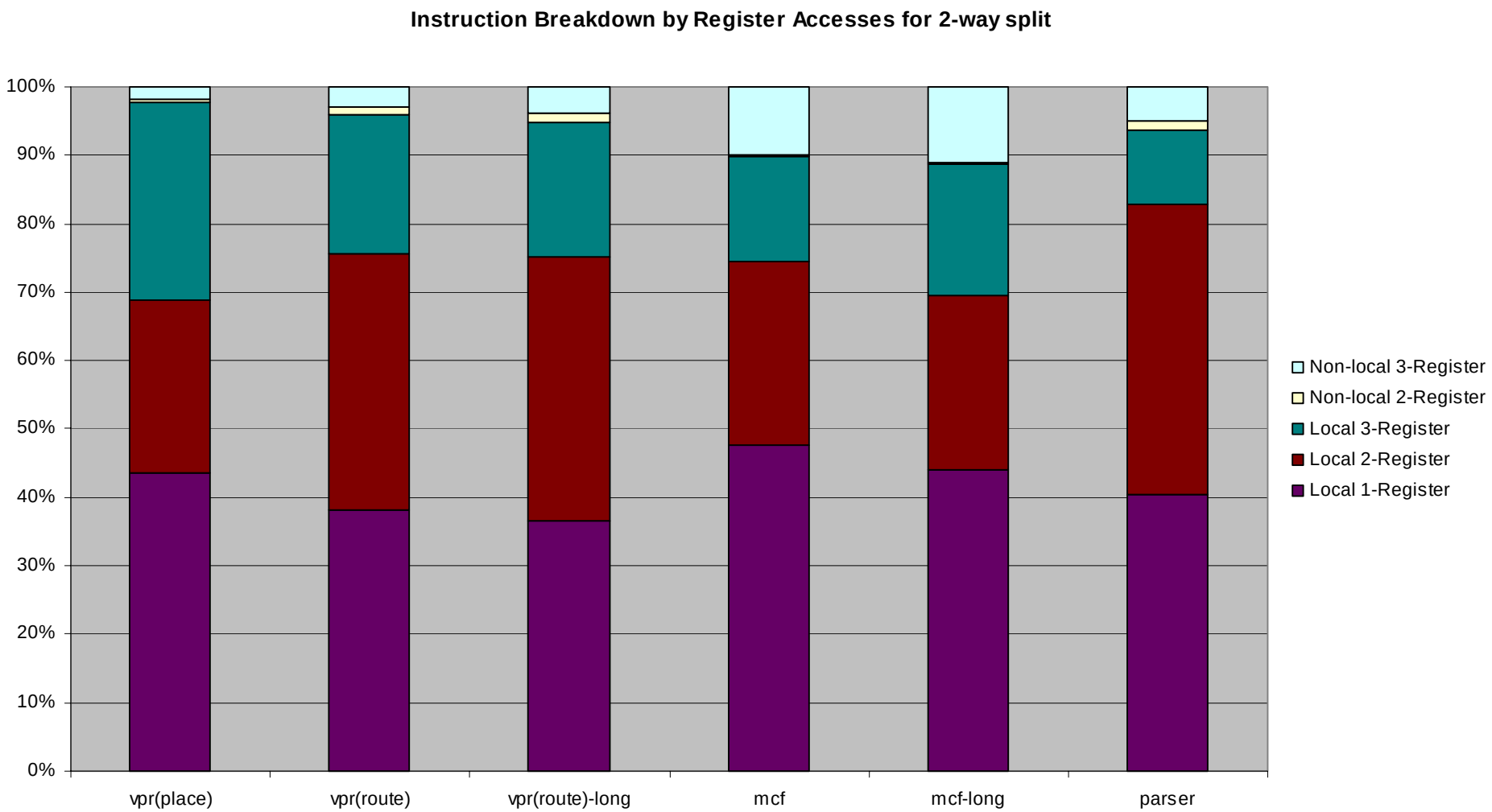
Register Files

- Register files will grow
  - Search for ILP
- Register access latencies will increase
  - Wire delay increases with register file size
  - Wire delay does not scale with technology

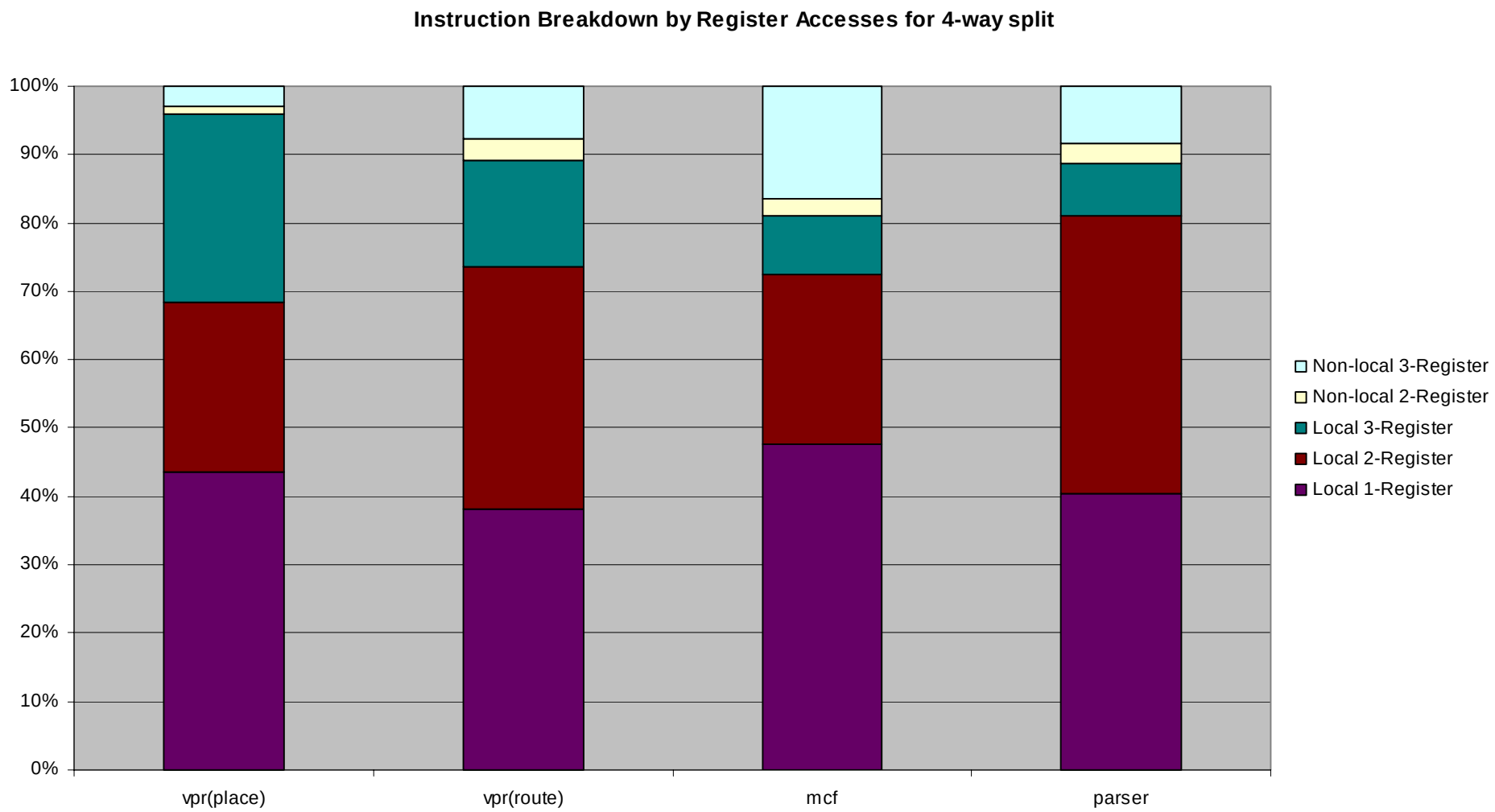
Hypothesis

- Splitting up the register file is a win
  - Reduced access time
  - Increased scalability
- But…
  - Inter-cluster delay
- The Hypothesis:
  - Dividing the register file and duplicating execution resources will reduce register latencies and improve processor performance.

Register Usage Patterns

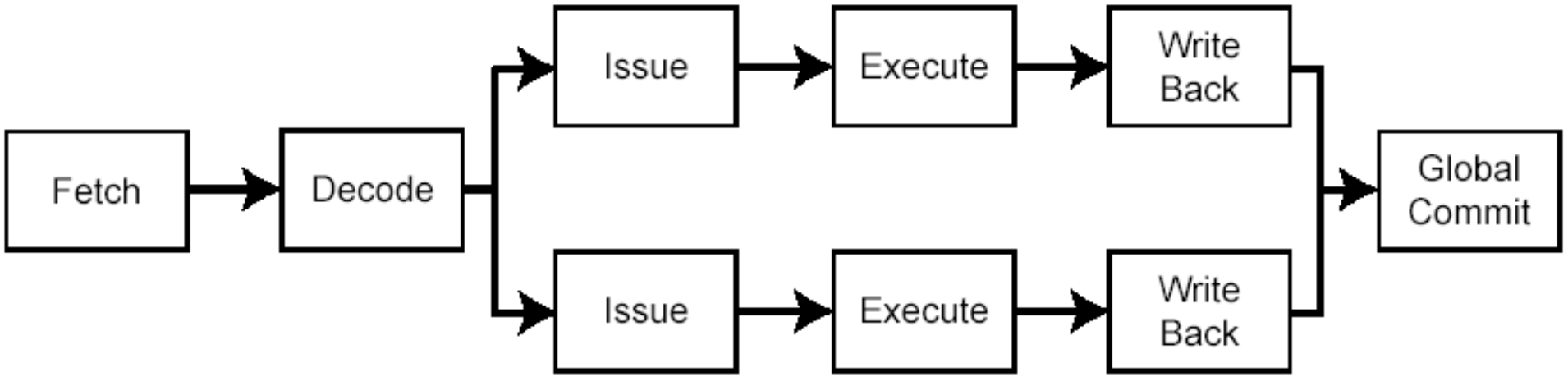


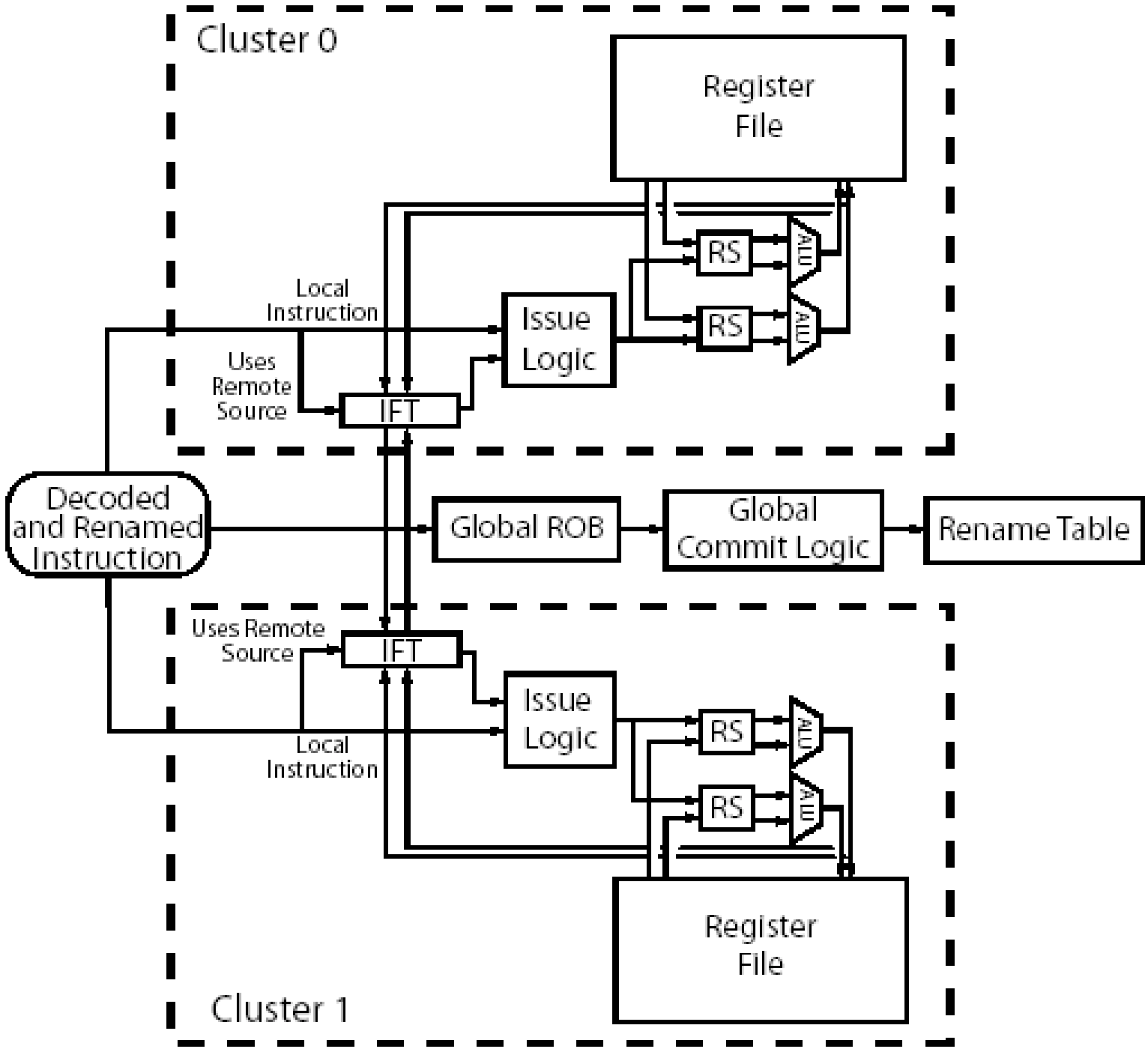
Register Usage Patterns



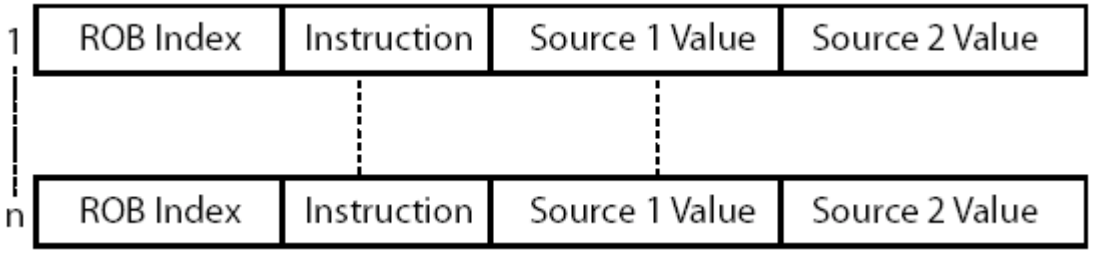
Pipeline

- Single instruction stream
- All clusters receive same instructions
  - Independently determine which should execute
- Global commit ensures in-order completion




Cluster 0
Register File
RS
RS
ALU
ALU
Local Instruction
Issue Logic
Uses Remote Source
IFT
Decoded and Renamed Instruction
Global ROB
Global Commit Logic
Rename Table
Uses Remote Source
IFT
Issue Logic
Local Instruction
RS
RS
ALU
ALU
Register File
Cluster 1

Instruction Forwarding Table(IFT)

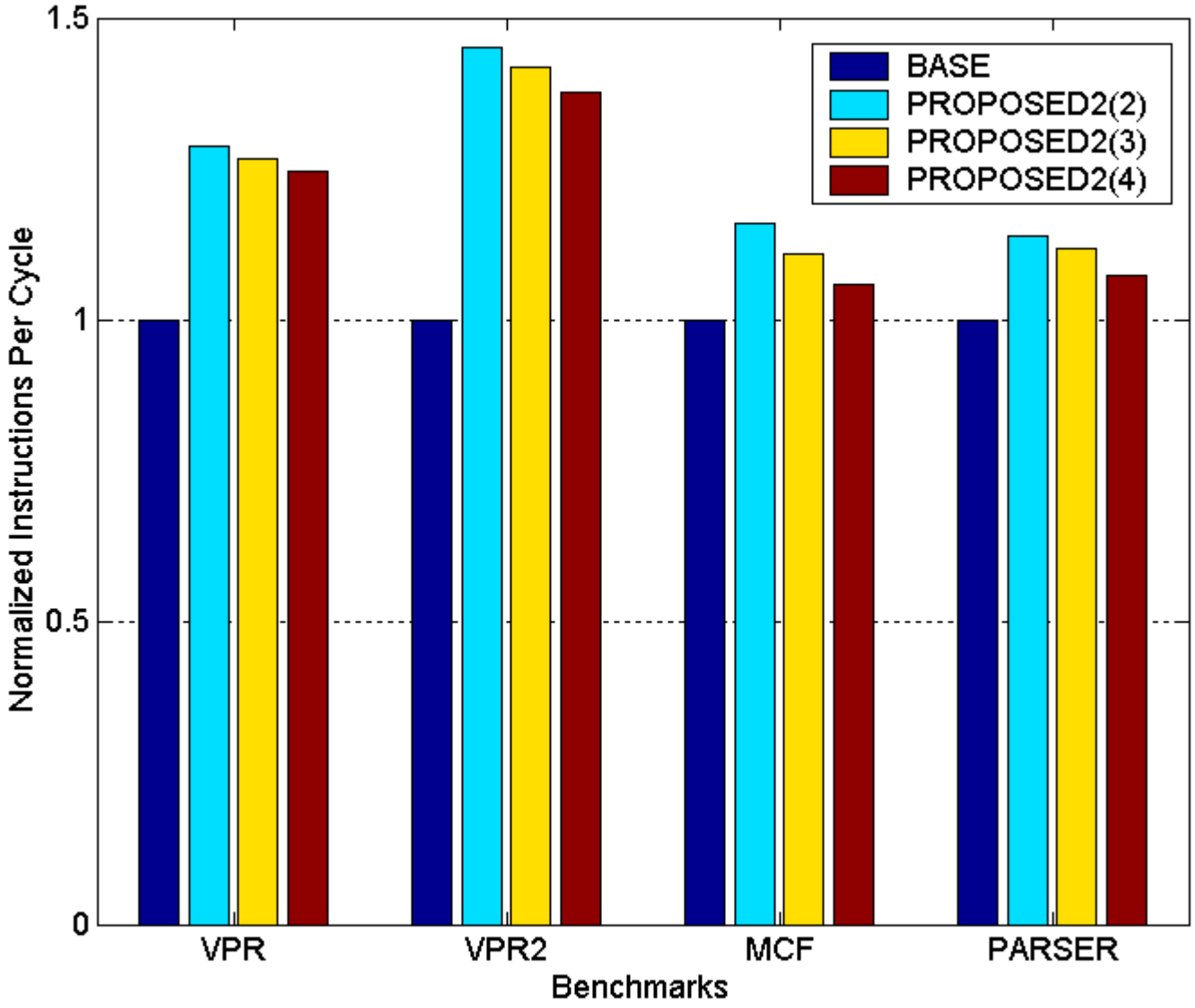


- If the correct data is ready, it forwards when the cluster receives the instruction
- Special table for tracking instructions that depend on results not yet calculated
  - Both sides allocate entries for such instructions
  - Sending side sends needed values when ready
  - Receiving side stores values until ready to issue

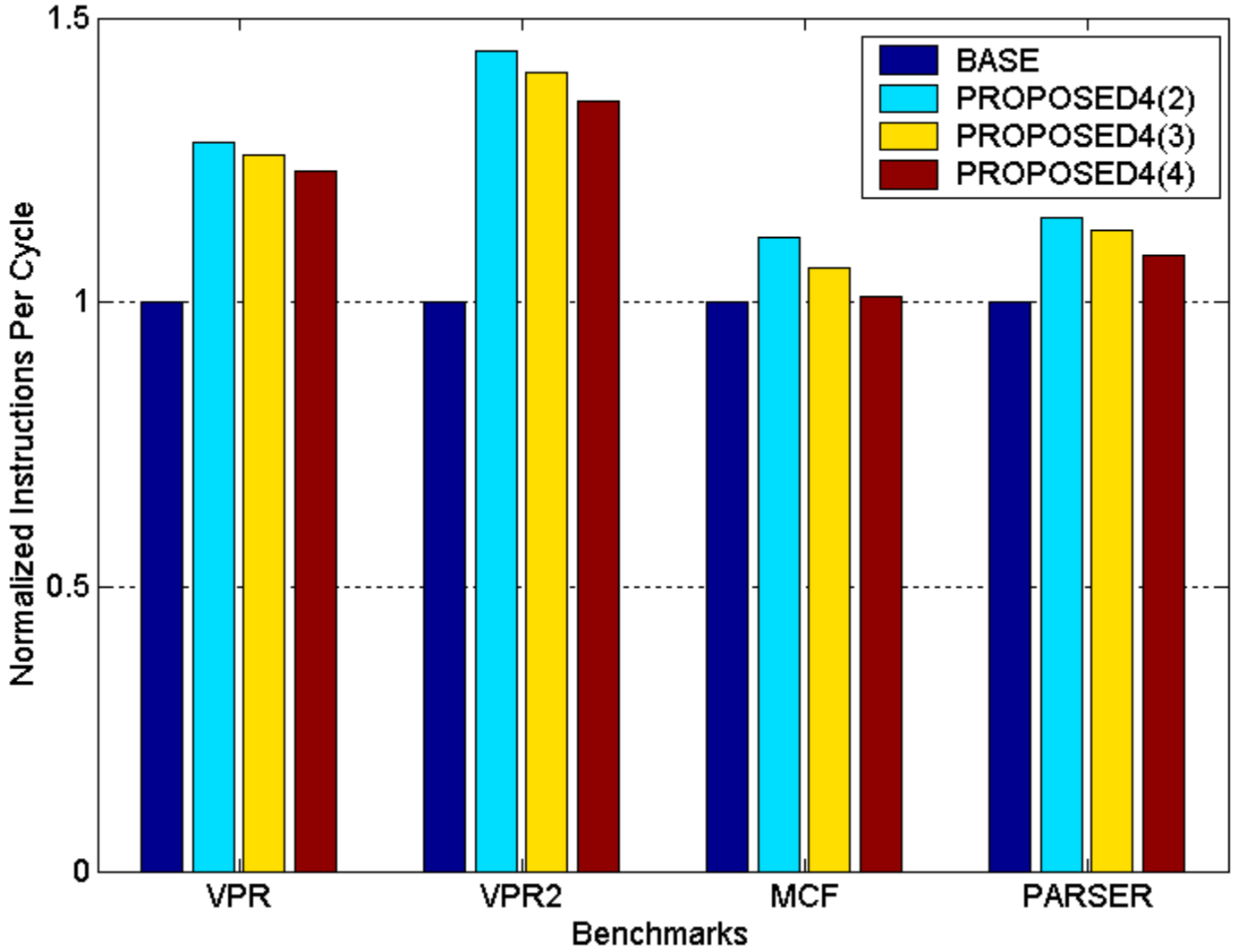
Experiment

- Modified SimpleScalar to simulate this architecture
- Baseline is our architecture with only one cluster and a two-cycle register file
- 4-wide issue processor in each cluster
- Simulated multi-cluster with 2 to 4 cycle delay for inter-cluster communication

2 Cluster Performance



4 Cluster Performance



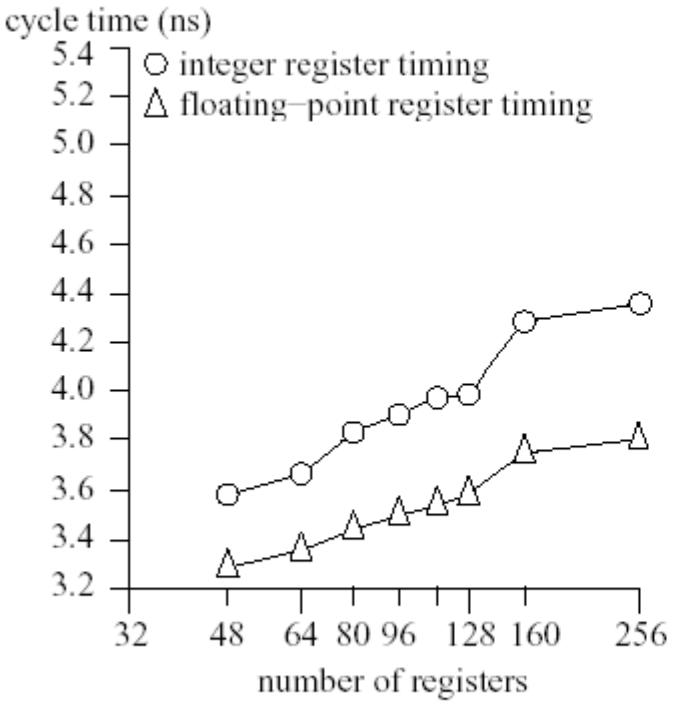
Conclusions

- Hypothesis is correct
  - Modest improvement

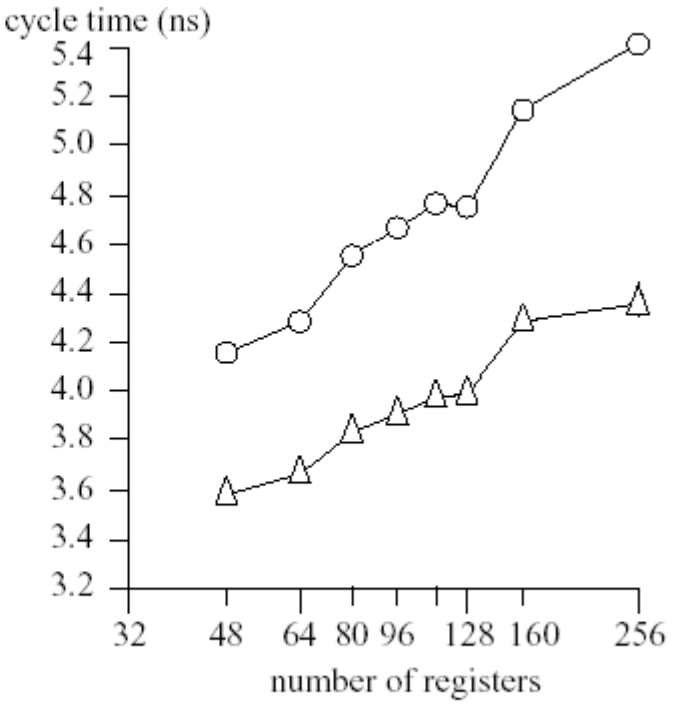
	2-cycle delay	4-cycle delay
2 clusters	14-45%	6-38%
4 clusters	11-44%	1-35%

- Scalable with current trends
  - Larger register files
  - Faster clocks and more emphasis on wire delay
  - Design localizes computation to reduce slowing effects of these trends
- Larger processors can use more clusters

Register Latencies



(a) Four–way issue processor



(b) Eight–way issue processor