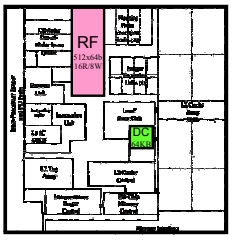
Banked Register File for SMT Processors

Jessica H. Tseng and Krste Asanović
MIT Computer Science and Artificial Intelligence Laboratory,
Cambridge, MA 02139, USA

BARC2004

Motivation

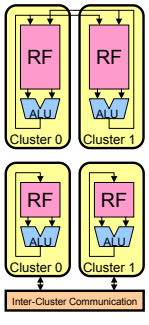
- Increasing demand on number of ports and number of registers in a register file.
- Growing concerns in access time, power, and die area.
  - Example: Alpha 21464 register file (RF) occupied over **5X** the area of 64KB primary data cache (DC).



Alpha 21464 Floorplan
ISSCC, 2002

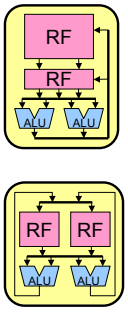
Distributed Architecture

- Duplicated
  - Fewer Read Ports
  - Same Number of Write Ports
  - Twice Total Number of Registers
  - Alpha 21264 & Alpha 21464
- Non-Duplicated
  - Fewer Read Ports
  - Fewer Write Ports
  - Complex Inter-Cluster Communication



Centralized Architecture

- Multi-Level: Register File Cache
  - Fewer Read Ports
  - Fewer Write Ports
  - Control Logic Complexity
  - Poor Locality
- One-Level Multi-Banked
  - Fewer Read Ports
  - Fewer Write Ports
  - Possible Conflicts
  - Control Logic Complexity
  - Possible Pipeline Stalls



Previous Work

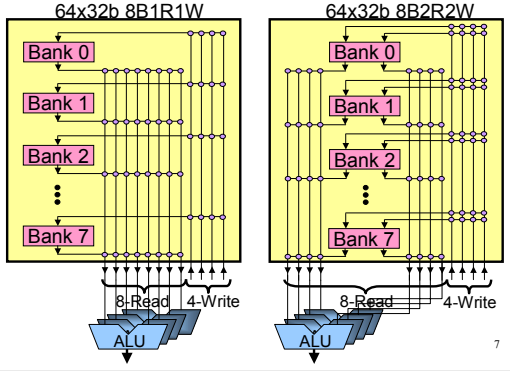
- Use minimal number of ports per register file banks: 1 or 2-read port(s) and 1-write port.
- Avoid issuing instructions that would cause register file read conflicts.
  - Add complexity to the critical wakeup-select loop for the issue logic → slower cycle time
- Resolve register file write conflicts by either delaying physical register allocation until write back stage or installing write buffers.
  - Complex pipeline control logic
  - Possible pipeline stalls

Our Work

- Use more ports per register file bank.
- Speculatively issue potentially conflicting instructions.
  - Minimize impact to the critical wakeup-select loop for the issue logic
- Rapidly repair pipeline and reissue conflicting instructions when conflicts are detected after issue.
  - No write buffer requirement
  - No pipeline stalls

Simpler and Faster Control Logic

Banked Regfile Structure

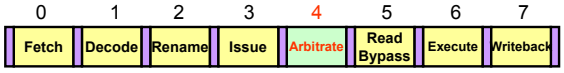


Regfile Characteristics

- Area: Magic, 0.25μm TSMC CMOS process
- Delay & Energy: HSPICE, 2.5V supply voltage

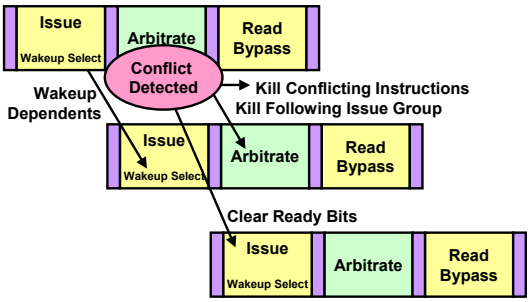
64x32b, 8 Read Ports & 4 Write Ports					
Type	Baseline	8B8R4W	8B2R2W	8B2R1W	8B1R1W
Area	100%	123%	37%	32%	30%
Delay	100%	83%	75%	75%	77%
Energy	100%	61%	59%	58%	41%
Packing Bitline					

Pipeline Structure



- Speculatively Issue Potentially Conflicting Instructions: Same Wakeup-Select Loop
- Additional Arbitration Pipeline Stage
  - Detect read and write bank conflicts when too many instructions try to read from or write to the same register file bank.
  - Mux operand addresses into available register file ports.
  - Adds a cycle to branch misprediction latency.

Pipeline Repair Operation



Performance Impact

- Additional Pipeline Stage
  - Increase in branch misprediction latency.
- Possibility of Bank Conflicts
  - Add penalty cycles to repair the pipeline.
  - Delay the issuing of dependent instructions.

Reduce Bank Conflicts

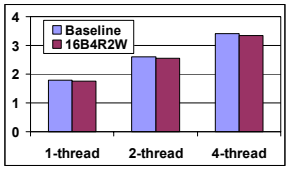
- Avoid contending for register file read ports when possible.
  - Bypass Skip: Operands that will be sourced from the bypass network do not compete for access to the register file.
  - Read Sharing: Allow multiple instructions to read the same physical register from the same bank.
- Help remove the correlation between accesses to the same bank.

Evaluating Performance Impact

- IPC degradation simulation: modify SMTSIM simulator to keep track of a unified physical register file organized into banks.
  - Register File Size: 512
  - Fetch, Dispatch, Commit Width: 8
  - Integer ALUs: 8
  - Memory Instructions: 4 (2-Load/2-Store)
- Benchmarks: Use SPEC CINT2000

IPC Comparison

- Average IPC degradation across SPEC CINT2000

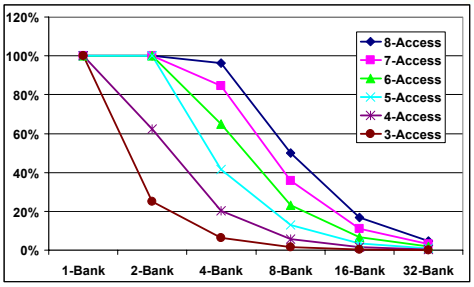


Workload	1-thread	2-thread	4-thread
Degradation	0.04	0.05	0.06
Degradation (%)	2.16	1.91	1.88
Range (%)	0.13~4.02	1.33~2.84	1.31~2.21

SMT Advantage

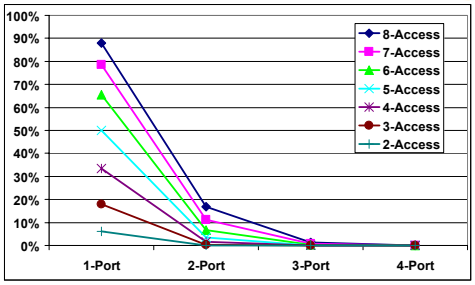
- Multi-threading Environment
  - Ability to hide the branch misprediction penalty.
- Large Number of Registers
  - More banks in the register file.
  - Dramatically reduce number of bank conflicts.

Bank Conflict Probability (1)



- Each bank has two local ports.

Bank Conflict Probability (2)



- Each design has sixteen banks.

Conclusion

- Banked register files work better for SMT processor than the single thread superscalar processor due to SMT's ability to hide branch misprediction penalty.
- Banked register files scale well with increasing issue widths in term of die area, access latency, and power.
  - Enough number of banks with enough number of ports.
  - The correlation of same-cycle accesses can be removed.
- For an eight-issue SMT processor with 512 physical registers, by adopting a 16-banked register file design with four read ports and two write ports, we can reduce regfile area by a factor of seven over a monolithic design while decreasing IPC by less than 2%.

Thank You