Advancing Near-Data Processing with Precise Exceptions and Effective Data Fetching

Sairo Santos, Tiago R. Kepe, Francis B. Moreira, Paulo C. Santos and Marco A. Z. Alves

Big Data

- Data-centric behavior;

  - Extensive data streaming;

  - Little data reuse;

 - Constantly accessing the main memory;

Near-Data Processing

- Memory wall;
    - Processing elements near the data;
    - A more data-centric system.

- Issues:
    - Added complexity;
    - System consistency;
    - Cache coherence;
   - Efficiency;

Near-Data Processing

- Strategies:
  - Limiting near-data action;
  - Strict in-order processing;
  - Forced inefficiency;
 - Especially when considering fine-grain architectures;

Vault

Logic layer

DRAM layers

Crossbar switch
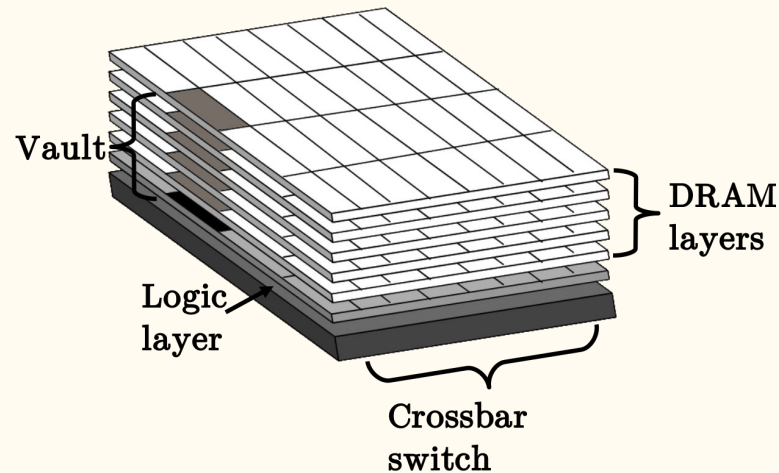
Proposed Modifications

- Two additional elements:
  - An instruction buffer;
    - Pooling of instructions;
    - No forced idleness;
    - Out or order operand loading;
    - Improved data throughput;
  - A memory disambiguation mechanism;
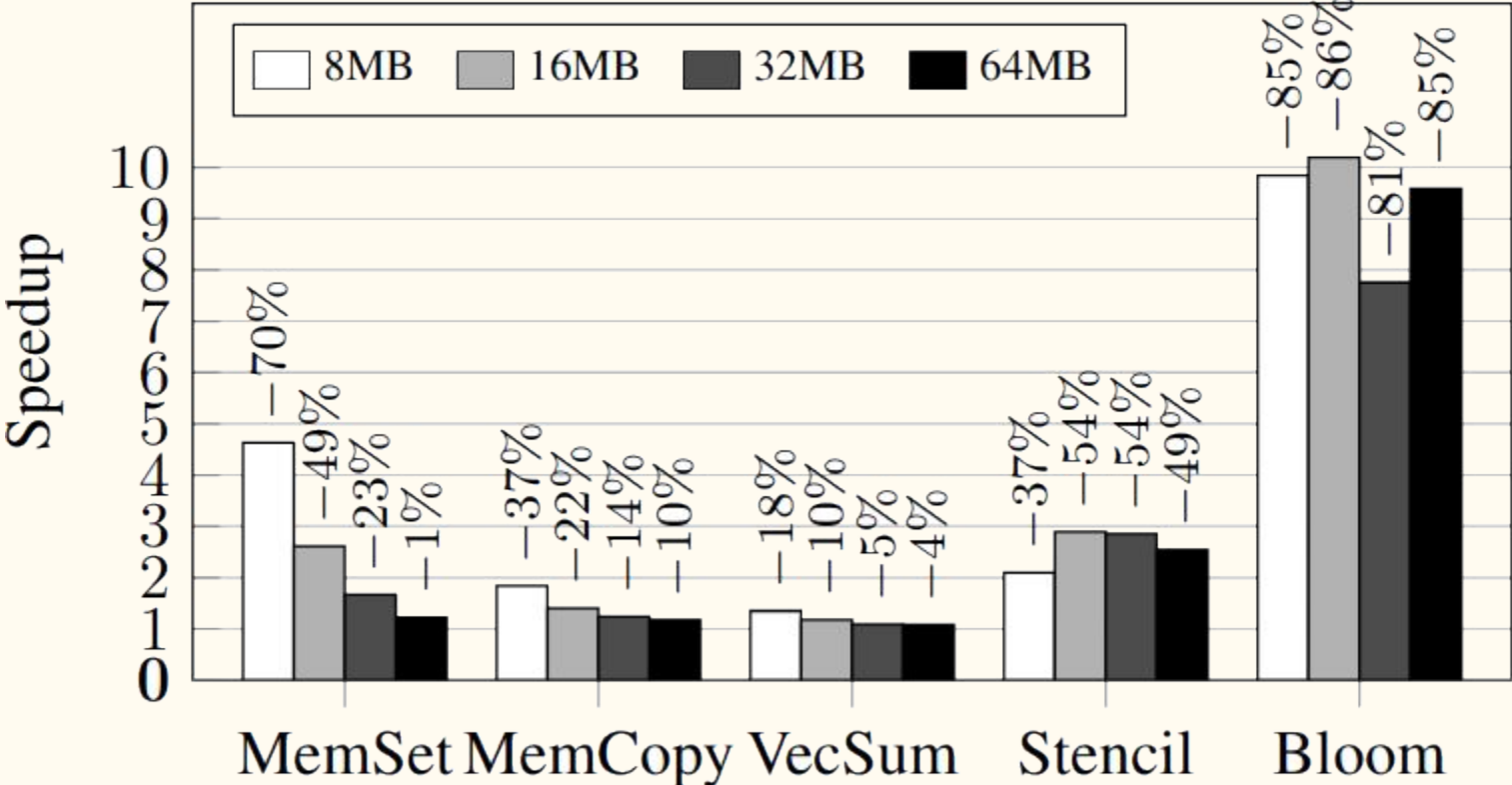   - Precise exceptions;

Vector-in-Memory Architeture
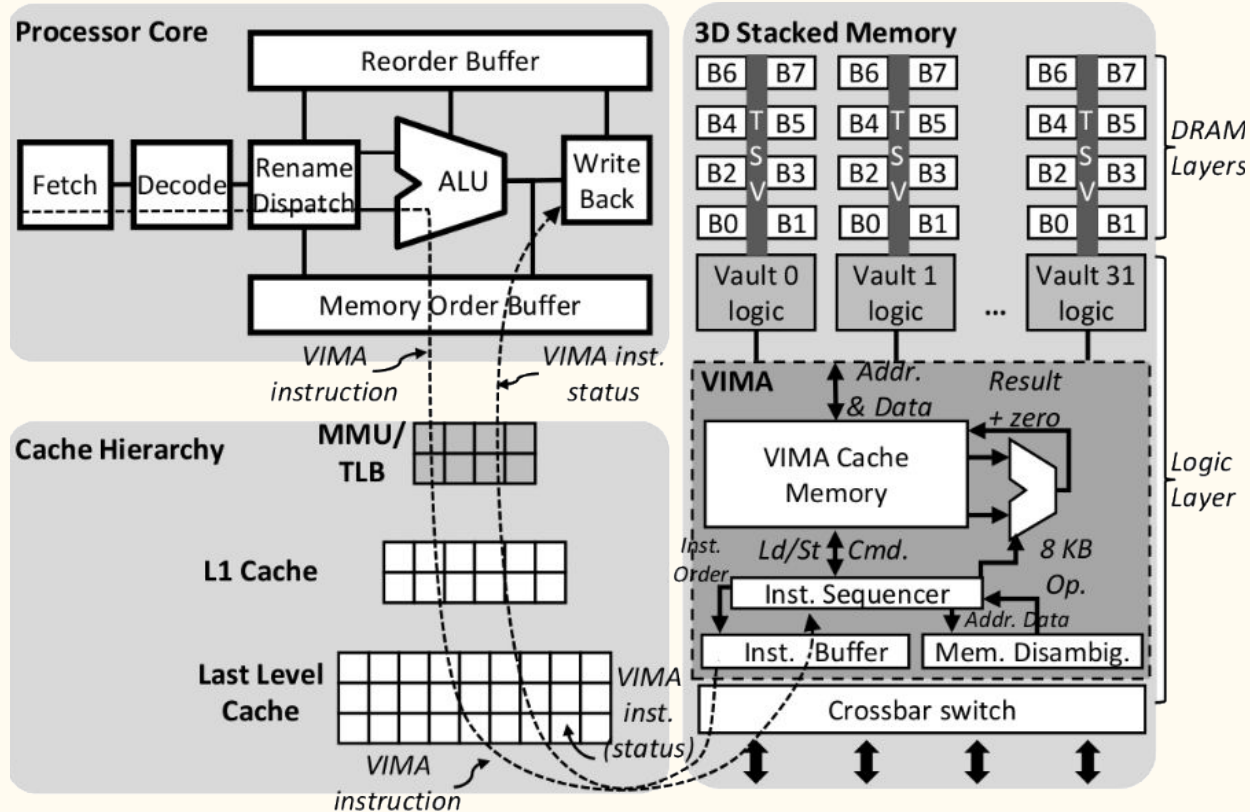
Results – Original Design
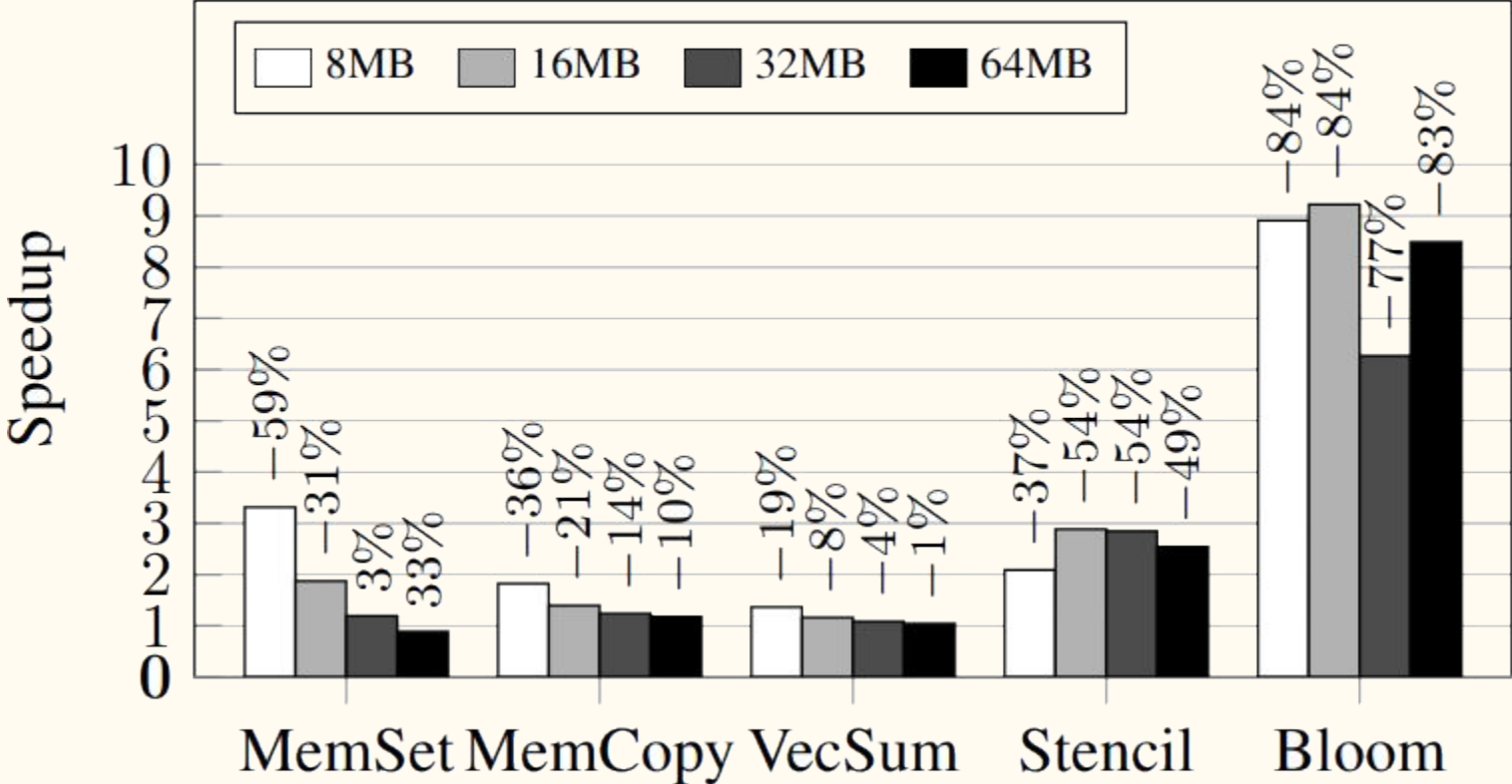
Results – Modified Design

Thank you!