Fast micro-architectural simulation on FPGAs using dynamic binary modification.

Oscar Palomar, John Mawer, Konstantinos Iordanou, Cosmin Gorgovan,
Andy Nisbet, Will Toms and Mikel Lujan
School of Computer Science, The University of Manchester
{first}.{last}@manchester.ac.uk

The University of Manchester

EPSRC
Engineering and Physical Sciences Research Council

1. New Architecture Challenges

- **Advances in run-times, compilers, microarchitecture, and chip fabrication?**
  - Delivering small improvements.
- **Increased design/implementation complexity?**
  - Power, performance, fault resilience/wearout issues.
  - Heterogeneity in microarchitectures, GPUs, DSPs.
  - Emergent workloads (computer vision, big-data).
- **HW/SW co-design as a potential solution – but it's hard, vertical expertise needed!**

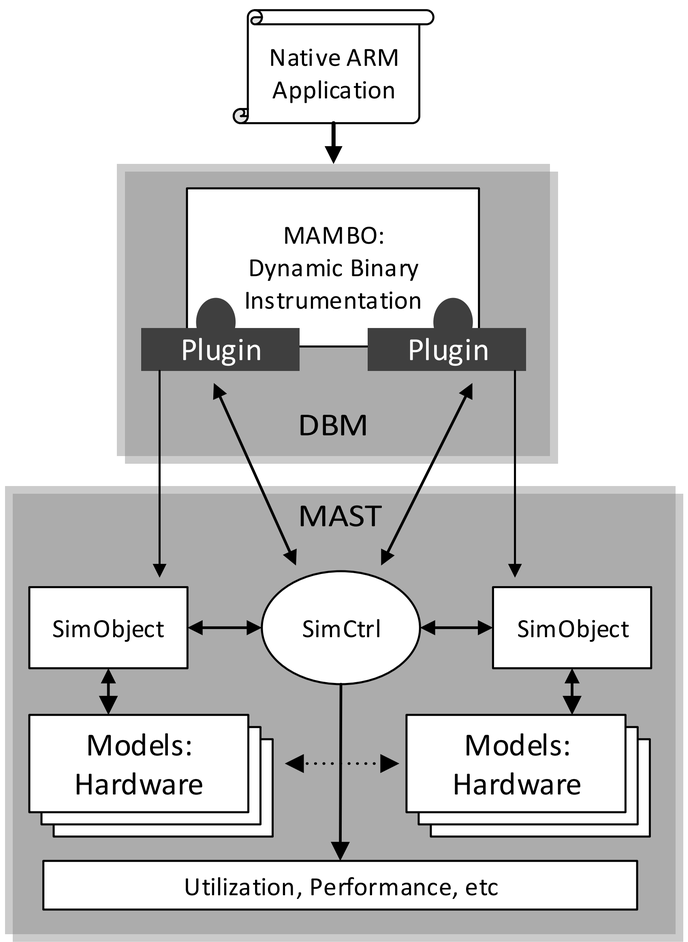
2. HW/SW Co-Design

- Creating an extensible co-designed infrastructure which:
  - **Efficiently evaluate μ-architecture IP.**
  - **Evaluate IP on real applications.**
  - **Directly use FPGA implementation of IP.**
  - **Easily support ISA extensions, custom hardware accelerators and entire processor cores.**

3. Mast Library

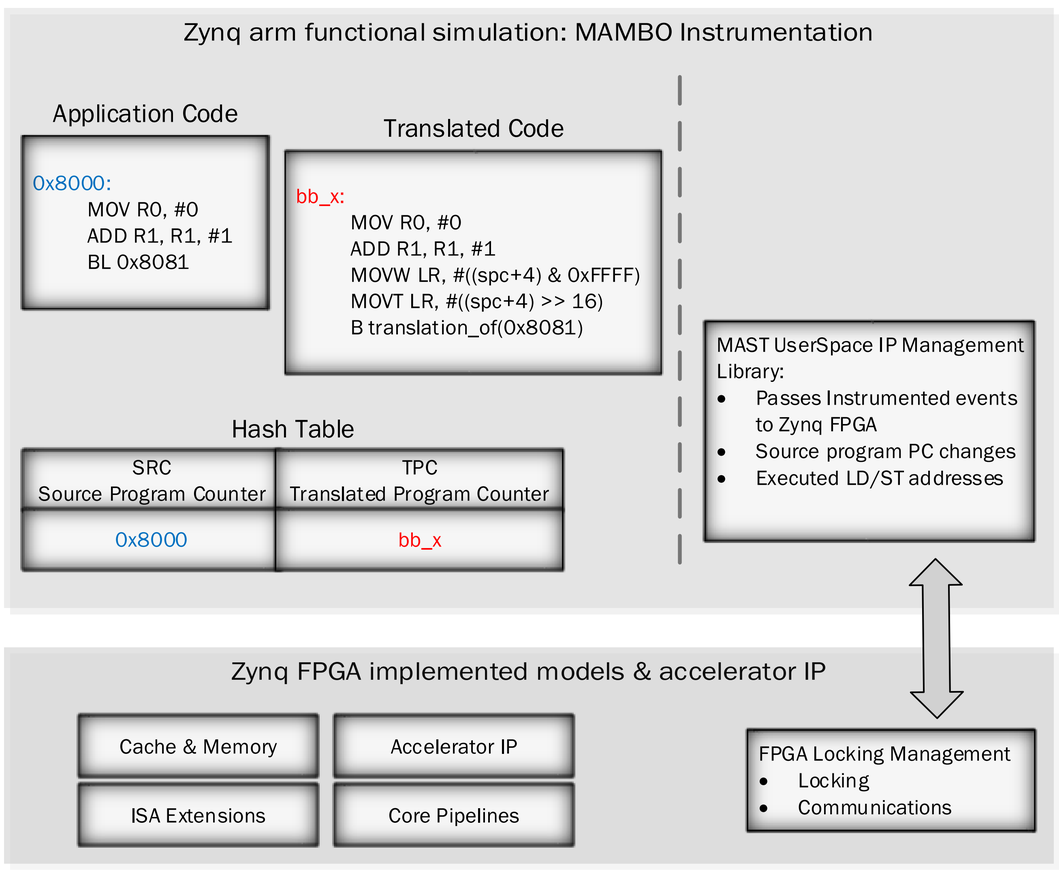
A C++ library and hardware interface standard/Bluespec library which:

- Discovers Hardware Blocks.
- Allows discovery of MAST compliant IP on an FPGA.
- Allows management of compliant hardware e.g locking blocks, reconfiguration of FPGA.
- Provides a userSpace interface to hardware.
- Enables IP accelerator blocks to be easily and efficiently integrated into applications.

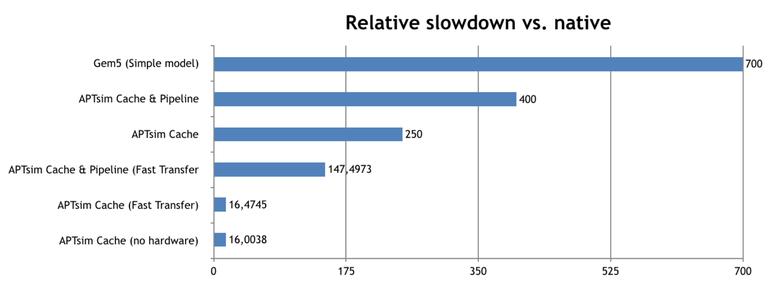


4. Simulator Design Overview

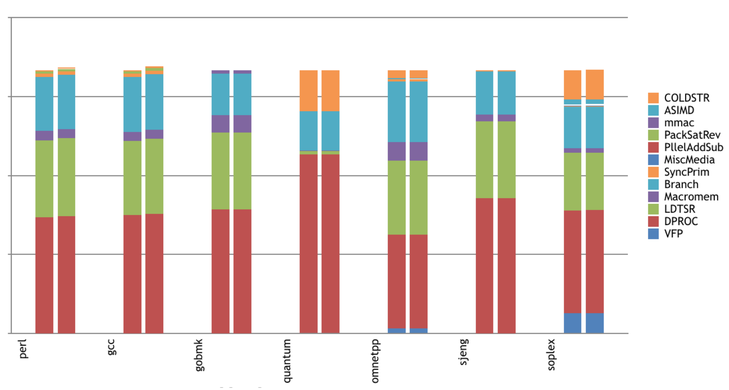
- MAMBO dynamic instrumentation **captures Load/store events including push/pop.**
- **PC changing events** branches, calls, returns.
- **Events passed to userspace IP management.**
- **Transferred to FPGA IP:**
  - CPU pipeline model (e.g CortexA7 in-order CPU)
  - Configurable coherent cache memory simulation.
  - Performance statistics for the pipeline model and the coherent cache System Hierarchy.



5. Performance and Validation

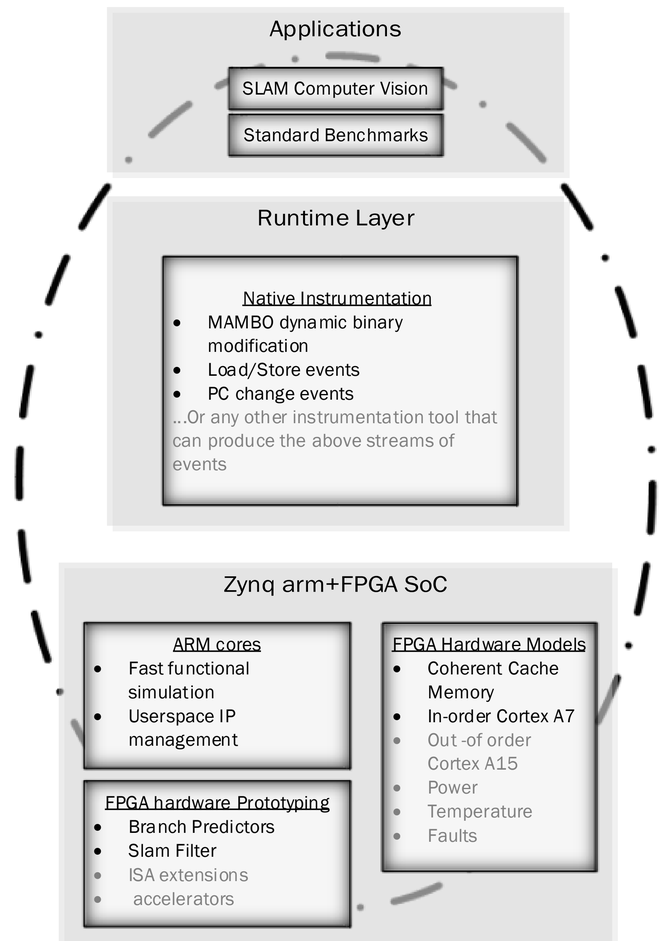


- Native and APTsim obtained with Zynq Cortex A9s@667MHz, 1G RAM
- GEM5 results obtained with Intel Xeon E3@3.2GHz with 32GB RAM



- Largest overall deviation is 1.5%

6. Future Work



7. References

[1] Cosmin Gorgovan, Amanieu d'Antras, and Mikel Luján. Mambo: A low-overhead dynamic binary modification tool for arm. *ACM Trans. Archit. Code Optim.*, 13(1):14:1–14:26, April 2016.

[2] J. Mawer, O. Palomar, C. Gorgovan, A. Nisbet, W. Toms, and M. Luján. The potential of dynamic binary modification and cpu-fpga socs for simulation. In *2017 IEEE 25th FCCM Annual International Symposium*, April 2017.