IMPLEMENTING A LEADING LOADS PERFORMANCE PREDICTOR ON COMMODITY PROCESSORS

BO SU[†] JOSEPH L. GREATHOUSE[‡] JUNLI GU[‡]
MICHAEL BOYER[‡] LI SHEN[†] ZHIYING WANG[†]

[†]NUDT [‡]AMD RESEARCH

JUNE 19, 2014

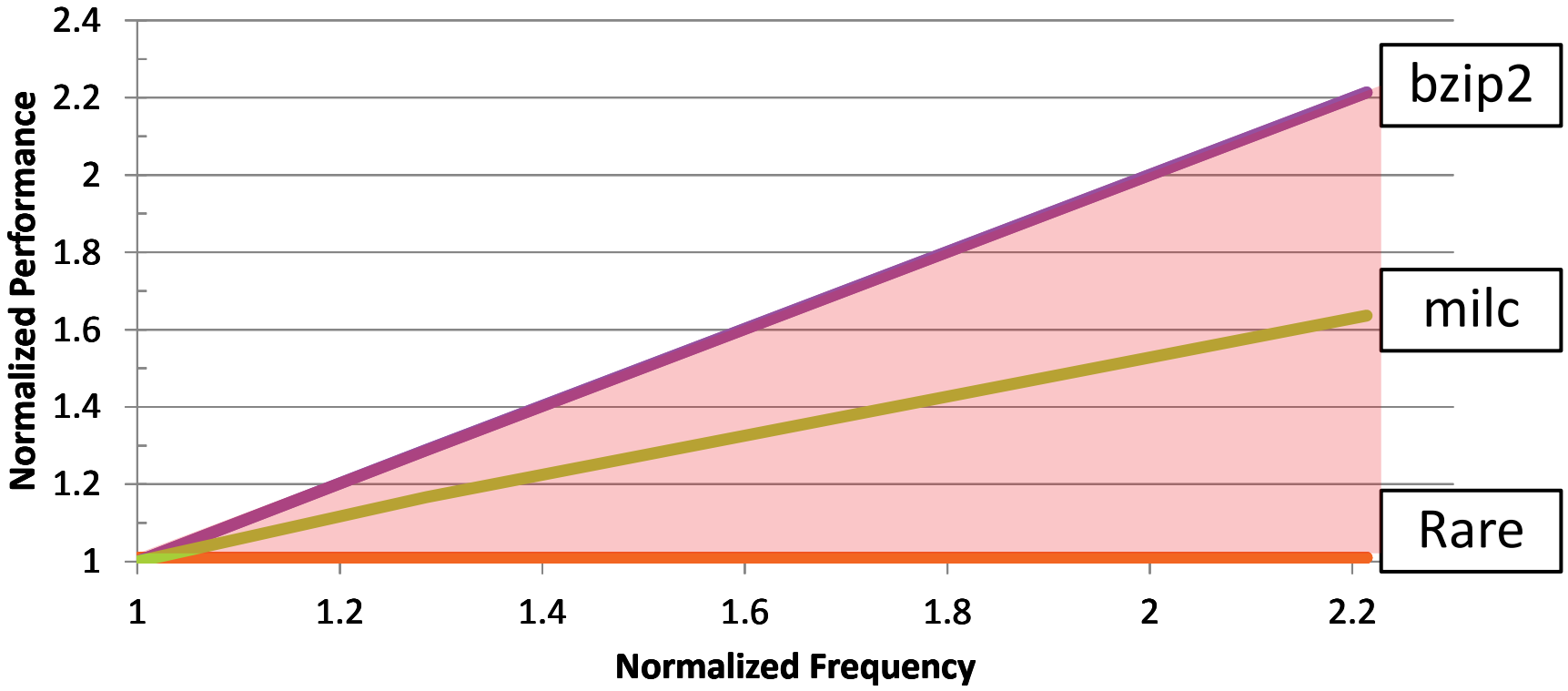
AMD

How fast is your application at different CPU frequencies?

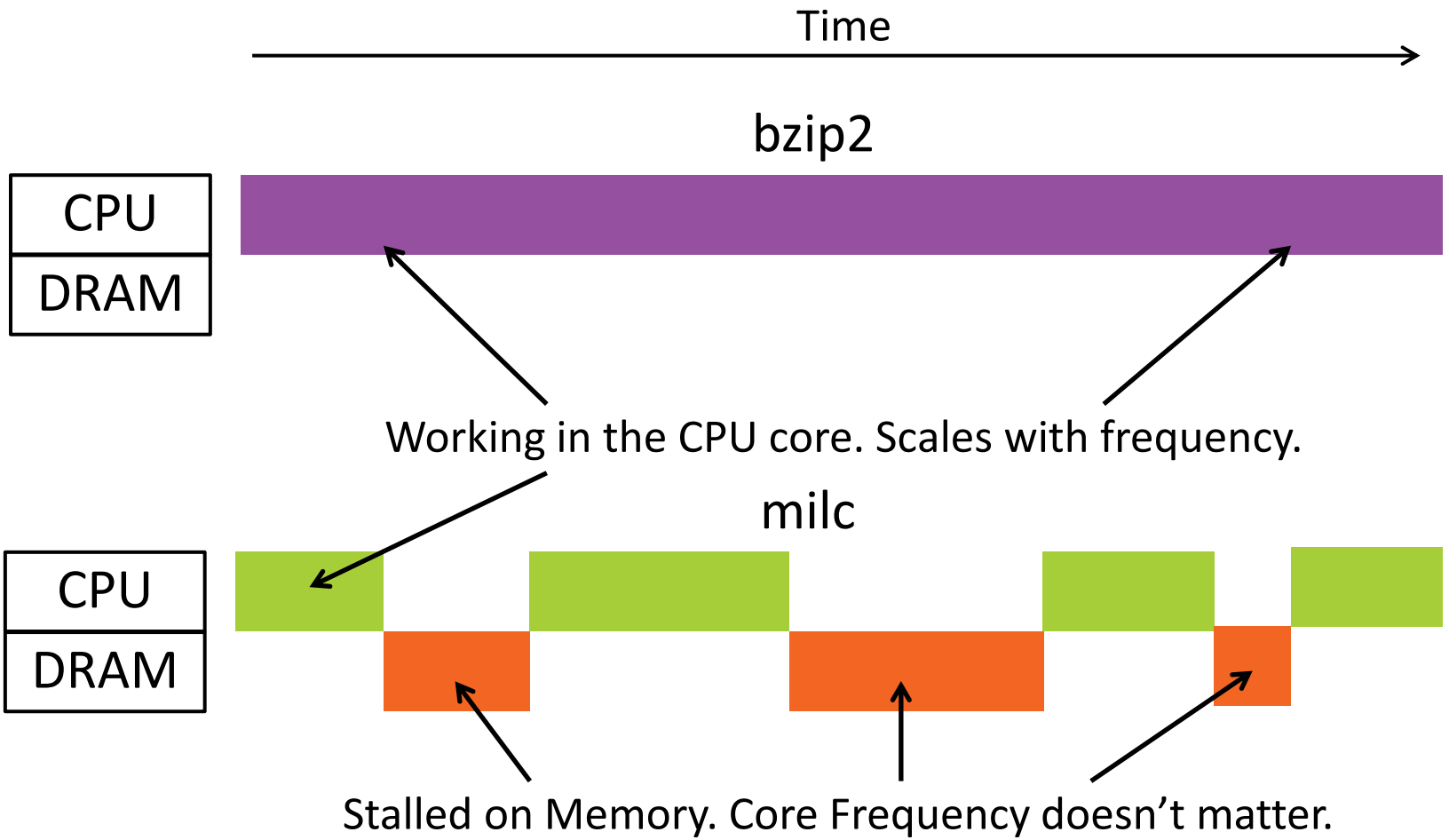
WHAT HAPPENS WHEN YOU CHANGE FREQUENCY?

AMD

Estimate #1	Nothing. Who cares about frequency?
Estimate #2	Performance difference is equal to frequency change.
Estimate #3	Something in between.



WHY DON'T NAÏVE ESTIMATES ALWAYS WORK?

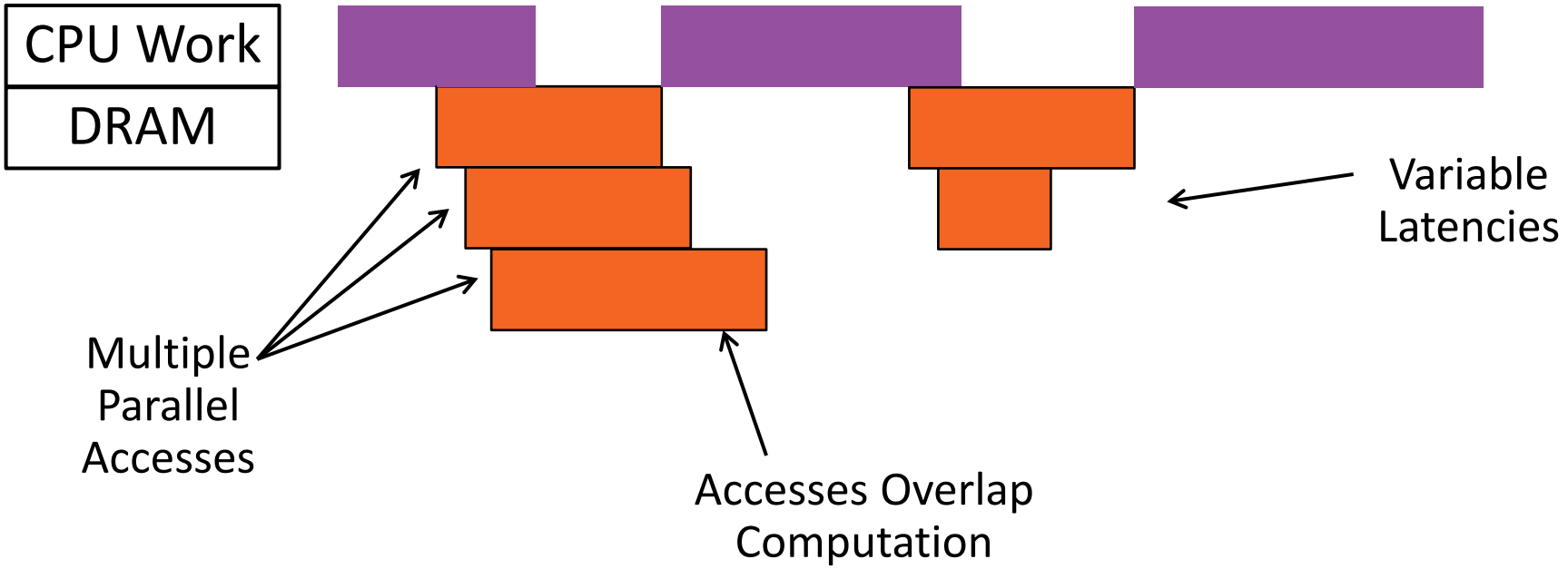


Core and memory time both matter

HOW DO YOU ESTIMATE “MEMORY TIME”?

MODERN CORES MAKE THIS DIFFICULT

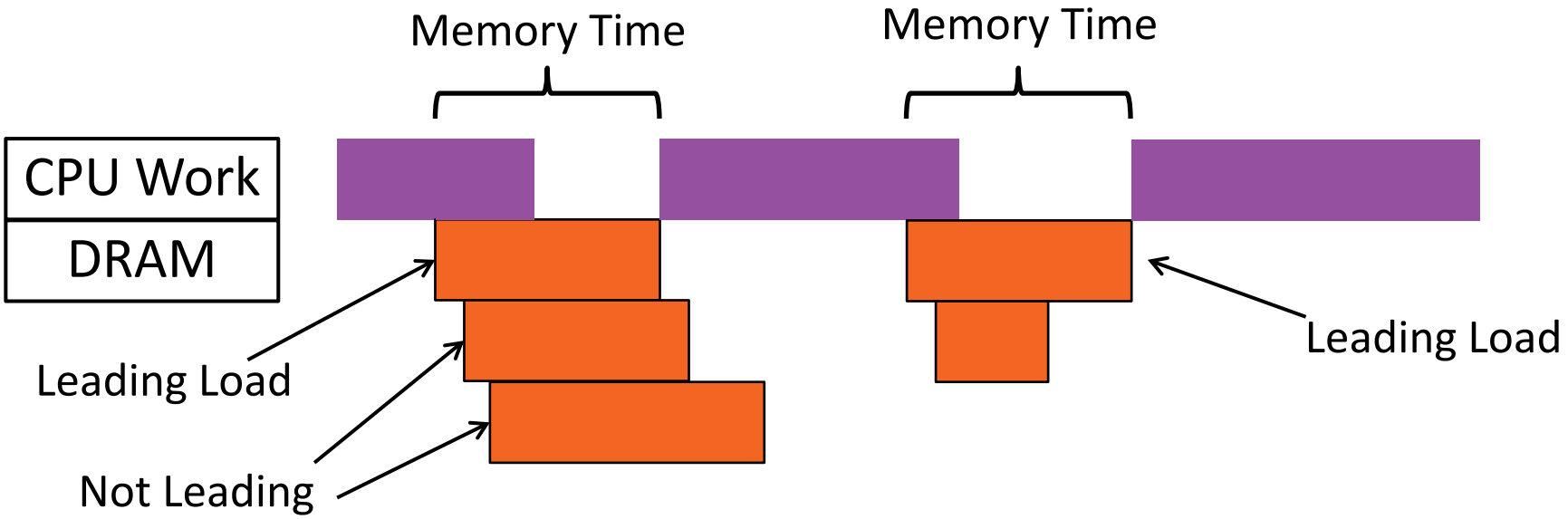
- Count the amount of time with an outstanding load?
- Count last-level cache misses?



“LEADING LOADS” MEMORY TIME ESTIMATION

AMD

Described by 3 separate groups in CF 2010, IEEE TOC, and IGCC 2011



Memory time approximately time that a leading loads is active

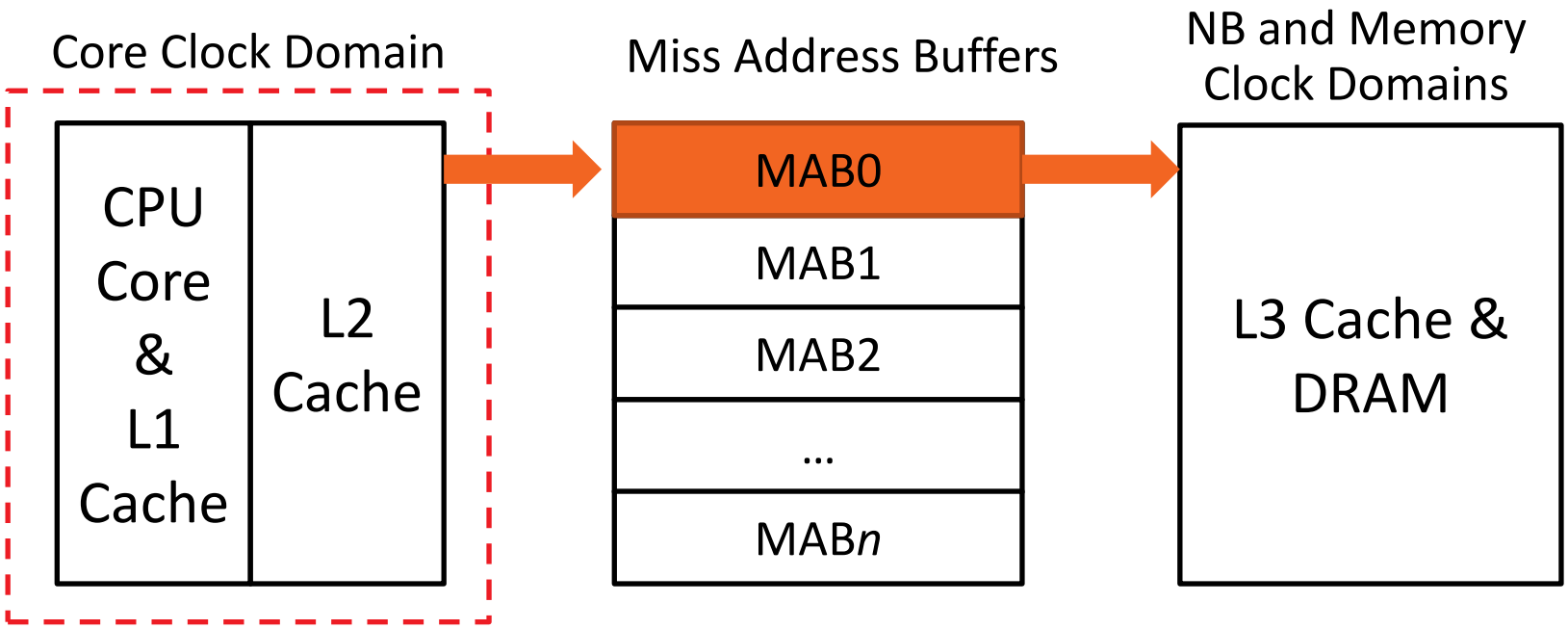
Simulation: ~0.2% estimation error across 2x change in frequency

LEADING LOADS ON AMD PROCESSORS

AMD

L2 cache misses held in Miss Address Buffer (MAB)

- MAB entries have a static priority (e.g. MAB0 is highest priority)
- Highest priority empty MAB holds the miss until it returns from memory



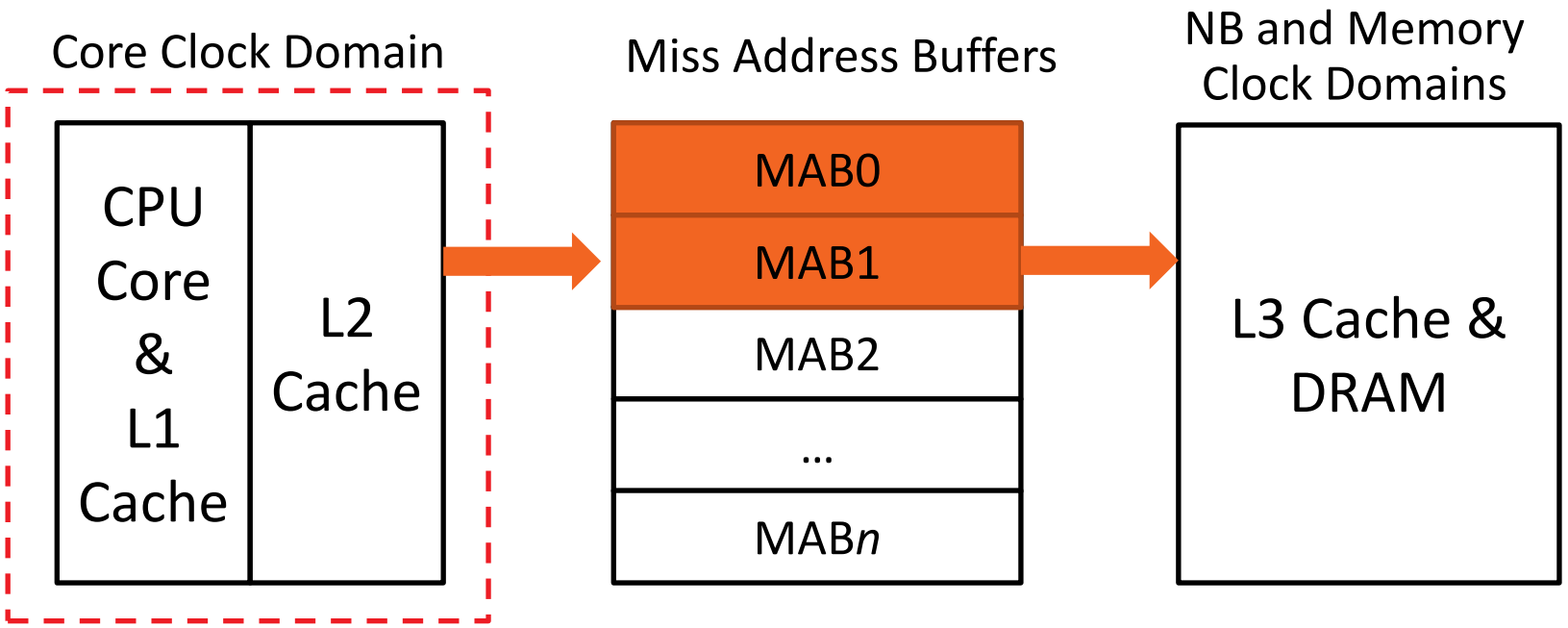
Performance event 0x69 allows SW to count # of cycles with filled MABs

LEADING LOADS ON AMD PROCESSORS

AMD

L2 cache misses held in Miss Address Buffer (MAB)

- MAB entries have a static priority (e.g. MAB0 is highest priority)
- Highest priority empty MAB holds the miss until it returns from memory



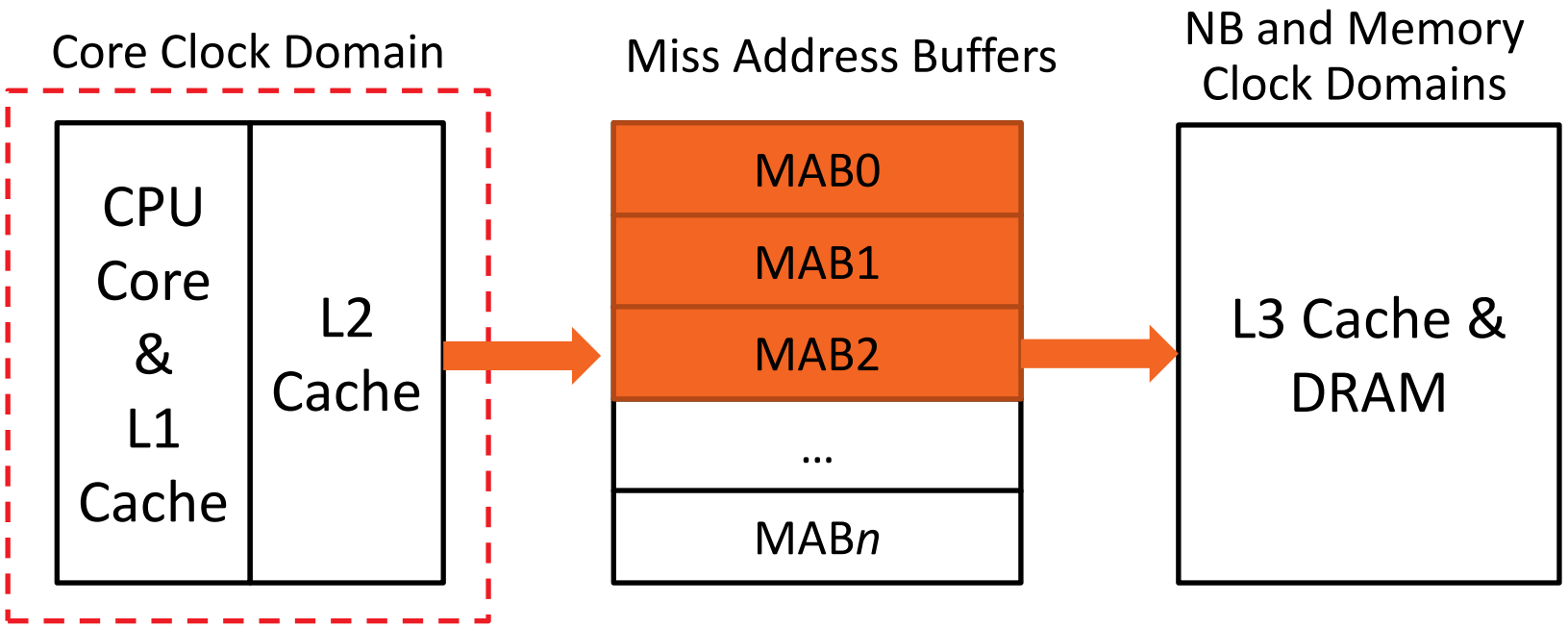
Performance event 0x69 allows SW to count # of cycles with filled MABs

LEADING LOADS ON AMD PROCESSORS

AMD

L2 cache misses held in Miss Address Buffer (MAB)

- MAB entries have a static priority (e.g. MAB0 is highest priority)
- Highest priority empty MAB holds the miss until it returns from memory



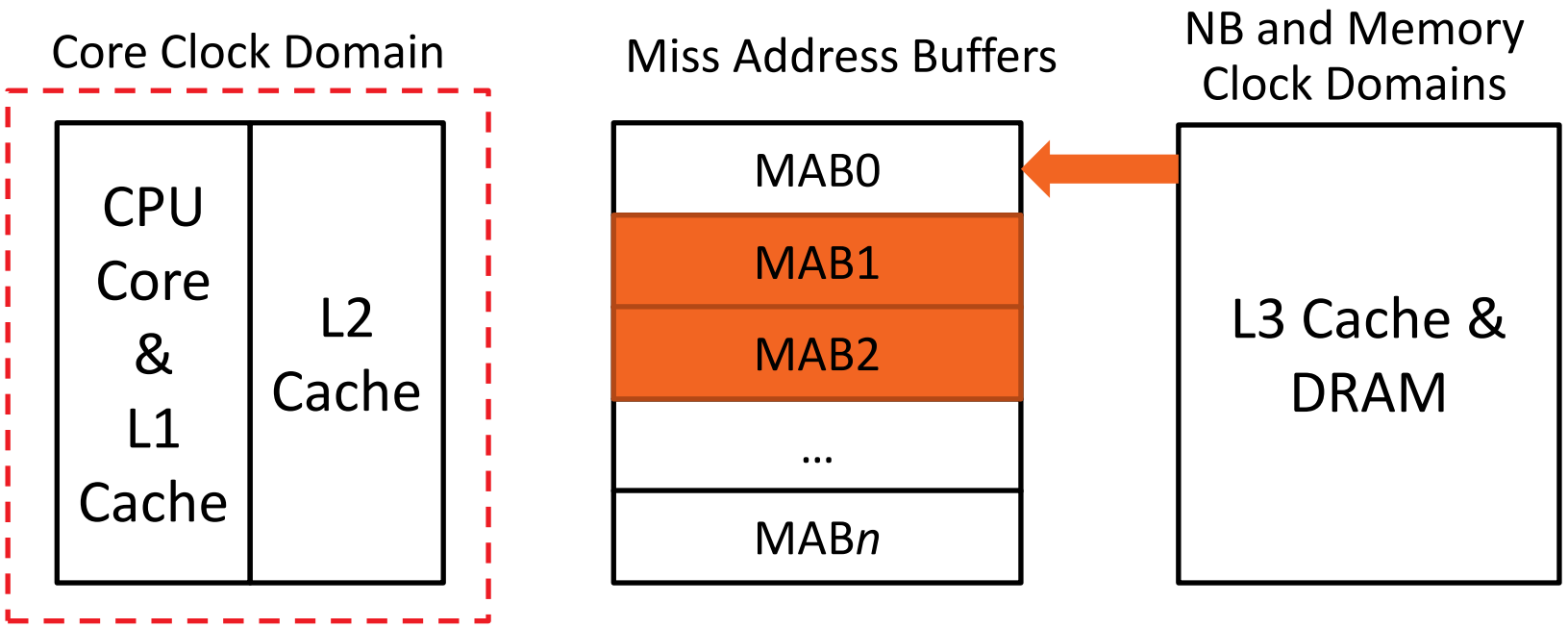
Performance event 0x69 allows SW to count # of cycles with filled MABs

LEADING LOADS ON AMD PROCESSORS

AMD

L2 cache misses held in Miss Address Buffer (MAB)

- MAB entries have a static priority (e.g. MAB0 is highest priority)
- Highest priority empty MAB holds the miss until it returns from memory



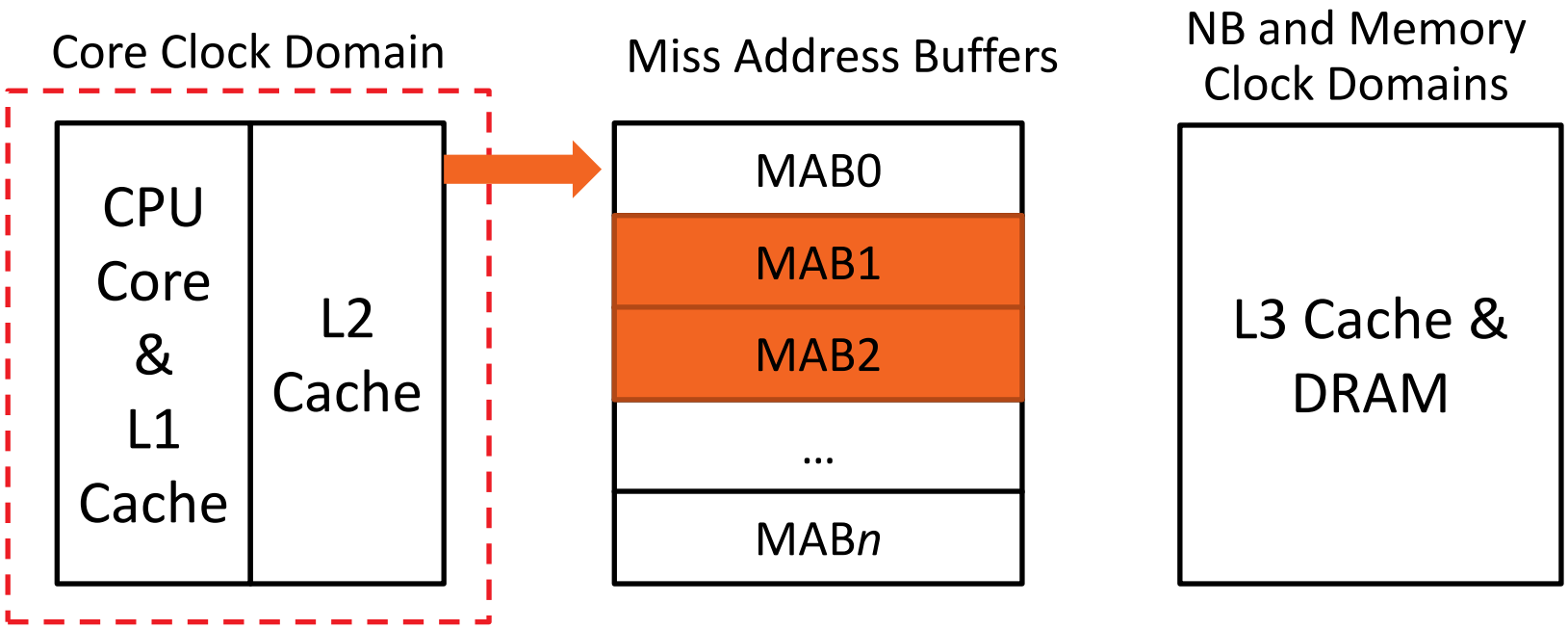
Performance event 0x69 allows SW to count # of cycles with filled MABs

LEADING LOADS ON AMD PROCESSORS

AMD

L2 cache misses held in Miss Address Buffer (MAB)

- MAB entries have a static priority (e.g. MAB0 is highest priority)
- Highest priority empty MAB holds the miss until it returns from memory



Performance event 0x69 allows SW to count # of cycles with filled MABs

PERFORMANCE ESTIMATION MODEL: LL-MAB

AMD

Measure occupancy time of the highest-priority MAB

- HW event 1: CPU Clocks not Halted (for Execution Time)
  - Performance Event 0x76
- HW event 2: MAB Wait Cycles (for Memory Time)
  - Performance Event 0x69
  - Family 15h Processors: Unit Mask 0

$$Memory\ Time(f1) = MAB\ Wait\ Cycles(f1)/(f1)$$

$$Core\ Time(f1) = Execution\ Time(f1) \text{-} Memory\ Time(f1)$$

$$Execution\ Time(f2) = Core\ Time(f1) * f1/f2 + Memory\ Time(f1)$$

EVALUATING PERFORMANCE PREDICTORS

AMD

- Run benchmarks at frequency 1, estimate runtime at frequency 2
- Run benchmark at frequency 2.
  - Difference between observed and estimated is **estimation error.**

- Estimation mechanisms:
  - Linear: Performance scales exactly with frequency (like bzip2)
  - Green Governor:
    - Count L3 cache misses
    - Assign delay to each cache miss
    - # Cache Misses * delay = “memory time”
  - LL-MAB: Count MAB0 cycles at “memory time”

EXPERIMENTAL SETUP

OTHER PROCESSORS TESTED IN THE PAPER

AMD

AMD Opteron™ 4386 Processor

- 2nd Generation Family 15h “Piledriver” CPU
- Minimum Frequency: 1.4 GHz, Maximum non-boost frequency: 3.1 GHz

Fedora® 19 Desktop (kernel version 3.10.6-200)

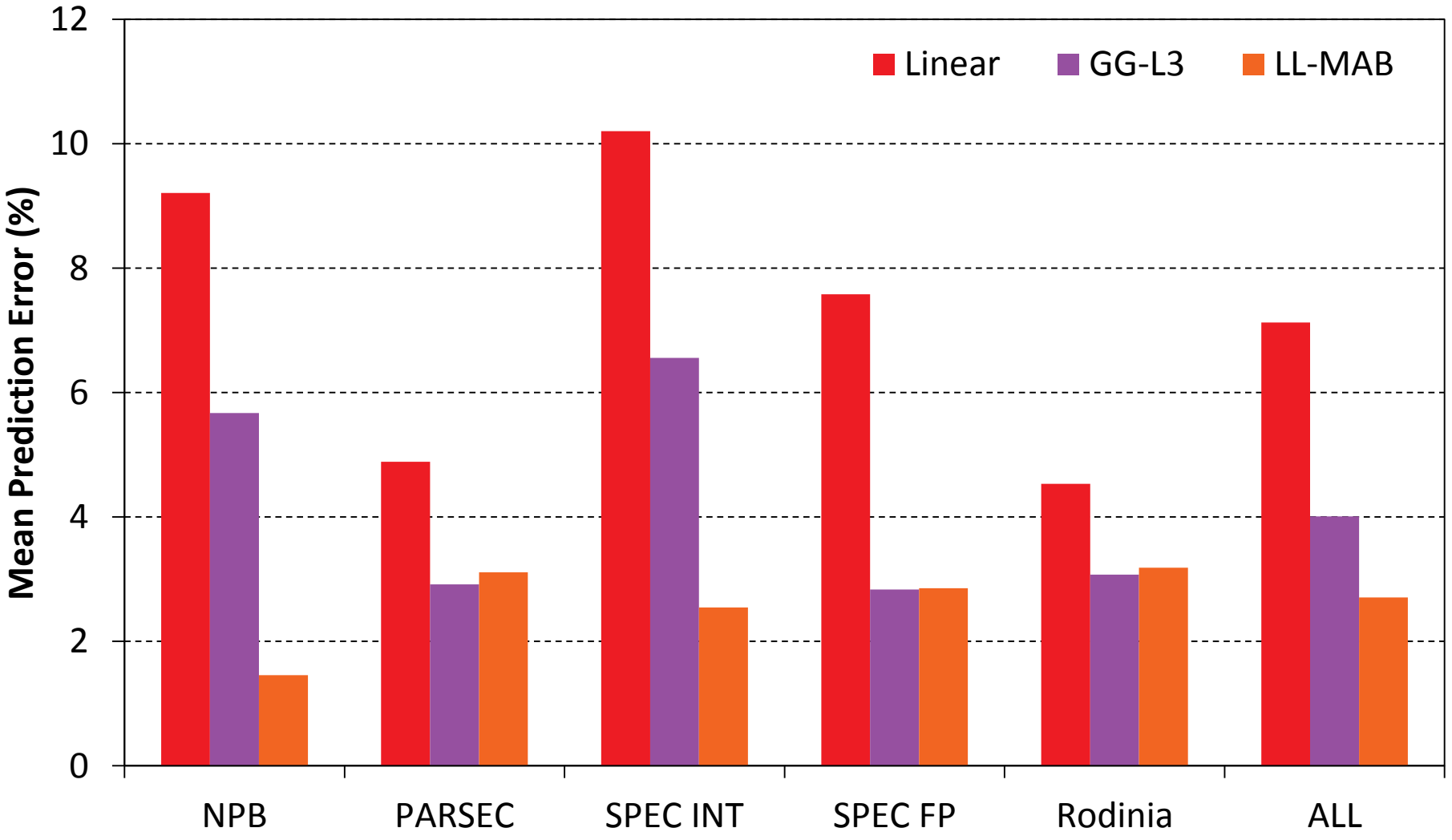
- Locked benchmarks to single core with *numactl*
- Used *msr-tools* to read performance counters around benchmark runs.
- *CPUFreq* userspace governor to manually control DVFS state. Boosting disabled.

66 Single-threaded benchmarks from:

- SPEC® CPU 2006
- NAS Parallel Benchmarks
- PARSEC
- Rodinia

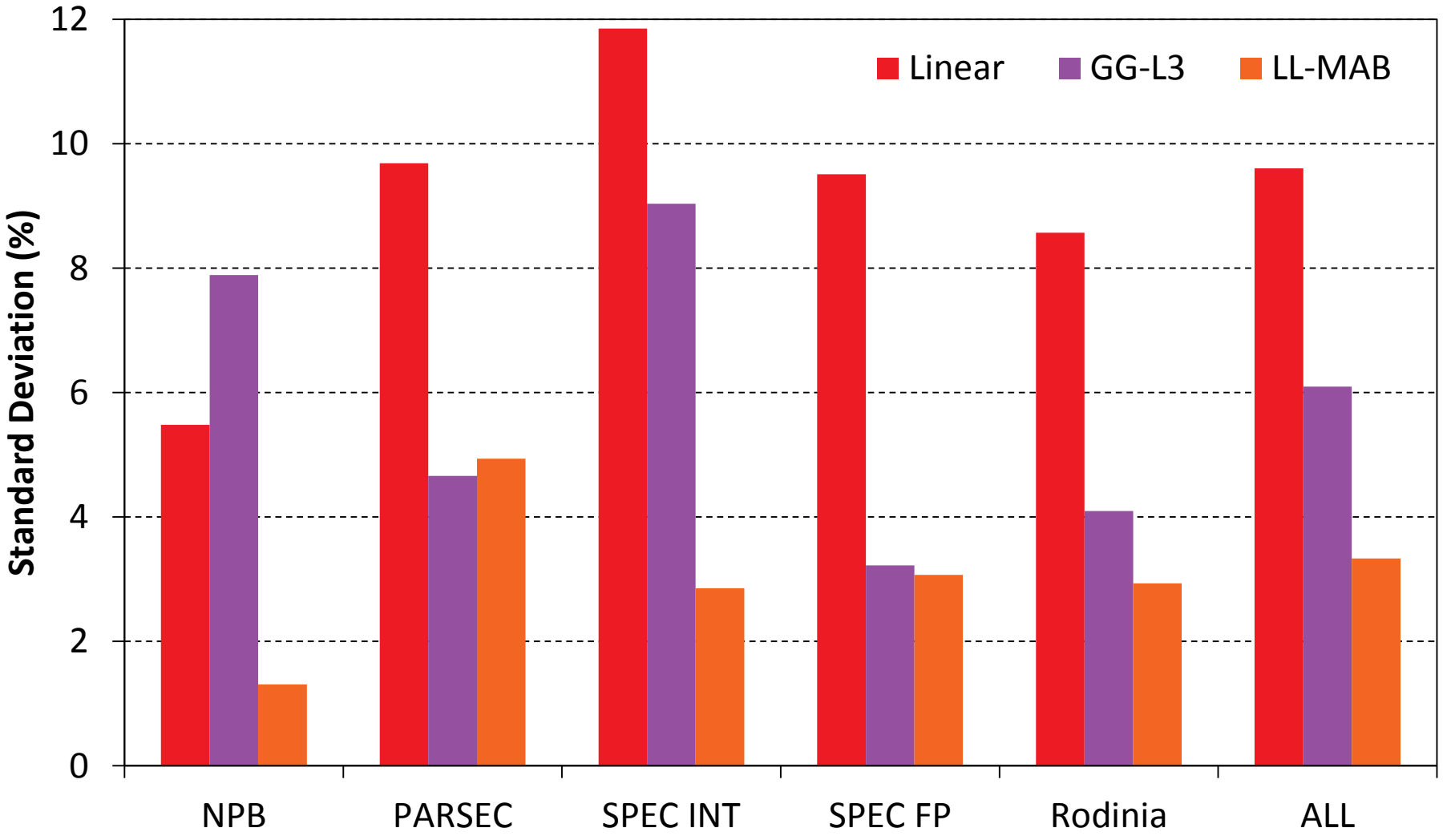
MEASURE AT 3.1 GHZ, ESTIMATE 1.4 GHZ RUNTIME

(LOWER IS BETTER)



STANDARD DEVIATION IS IMPORTANT FOR PREDICTIONS

(LOWER IS STILL BETTER)



DISCLAIMER & ATTRIBUTION

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions and typographical errors.

The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product releases, product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. AMD assumes no obligation to update or otherwise correct or revise this information. However, AMD reserves the right to revise this information and to make changes from time to time to the content hereof without obligation of AMD to notify any person of such revisions or changes.

AMD MAKES NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE CONTENTS HEREOF AND ASSUMES NO RESPONSIBILITY FOR ANY INACCURACIES, ERRORS OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION.

AMD SPECIFICALLY DISCLAIMS ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE. IN NO EVENT WILL AMD BE LIABLE TO ANY PERSON FOR ANY DIRECT, INDIRECT, SPECIAL OR OTHER CONSEQUENTIAL DAMAGES ARISING FROM THE USE OF ANY INFORMATION CONTAINED HEREIN, EVEN IF AMD IS EXPRESSLY ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

ATTRIBUTION

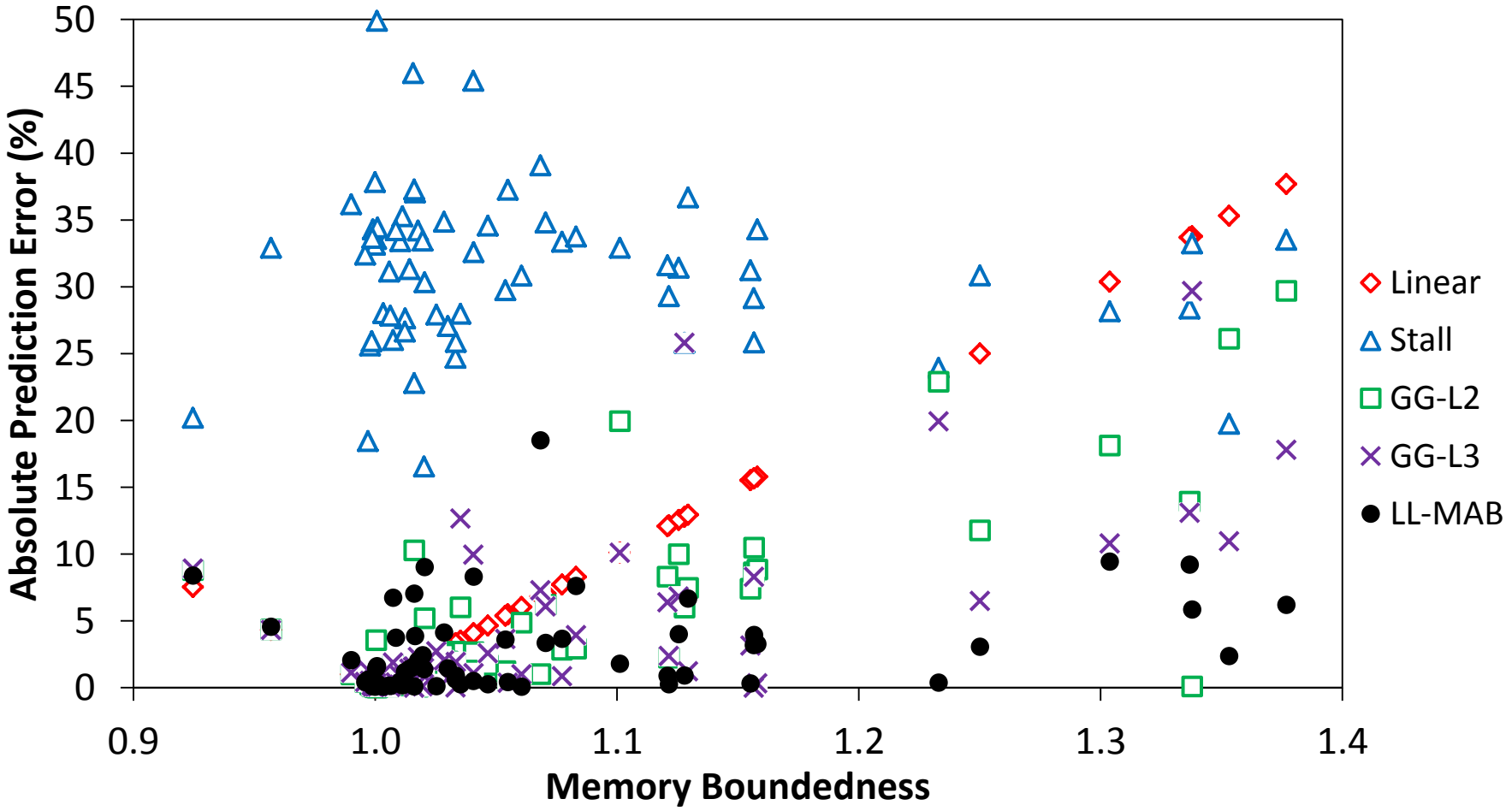
Backup Slides

PREDICTION ACCURACY PER BENCHMARK

AMD

Memory Boundedness = Ratio of execution cycles at two frequencies

– 1.0 = no change in cycles (completely compute bound, e.g. bzip2)



CONCLUSION

AMD

- First leading loads implementation on real processors
- Higher accuracy than existing predictors
- Lower accuracy than simulation due to HW complexity
- Lightweight estimation mechanism (only requires 2 counters)
  - Path to better performance and power prediction