Understanding native event semantics

9th JLESC Workshop

Anthony Danalis, Heike Jagode, Daniel Barry, Jack Dongarra

Knoxville, TN

Apr 16th, 2019

INNOVATIVE COMPUTING LABORATORY
THE UNIVERSITY of TENNESSEE

Performance API (PAPI): Overview

http://icl.cs.utk.edu/papi/

Latest PAPI Version 5.7

INNOVATIVE COMPUTING LABORATORY
THE UNIVERSITY of TENNESSEE

PAPI

- Library that provides a **consistent interface** (and methodology) for hardware performance counters, found across the system:
  i. e., CPUs, GPUs, on-/off-chip Memory, Interconnects, I/O system, File System, Energy/Power, etc.
- PAPI enables software engineers to see, in near real time, the **relation** between
  SW performance and **HW events across the entire compute system**

PAPI

- Library that provides a **consistent interface** (and methodology) for hardware performance counters, found across the system:
  i. e., CPUs, GPUs, on-/off-chip Memory, Interconnects, I/O system, File System, Energy/Power, etc.
- PAPI enables software engineers to see, in near real time, the relation between **SW performance** and **HW events across the entire compute system**

SUPPORTED ARCHITECTURES:

- AMD
- IBM Blue Gene Series
- IBM Power Series
- Intel Westmere, Sandy|Ivy Bridge, Haswell, **Broadwell**, **Skylake**, KNC, **Knights Landing**

PAPI

- Library that provides a **consistent interface** (and methodology) for hardware performance counters, found across the system:
  i. e., CPUs, GPUs, on-/off-chip Memory, Interconnects, I/O system, File System, Energy/Power, etc.
- PAPI enables software engineers to see, in near real time, the relation between **SW performance** and **HW events across the entire compute system**

SUPPORTED ARCHITECTURES:

- AMD
- ARM Cortex A8, A9, A15, ARM64
- CRAY: Gemini and Aries interconnects, power
- IBM Blue Gene Series, Q: 5D-Torus, I/O system, EMON power/energy
- IBM Power Series
- Intel Westmere, Sandy|Ivy Bridge, Haswell, **Broadwell**, **Skylake**, KNC, **Knights Landing**
- Intel KNC, **Knights Landing power/energy**
- Intel RAPL (power/energy); **power capping**
- InfiniBand
- Lustre FS
- NVIDIA Tesla, Kepler: CUDA support for multiple GPUs; PC Sampling
- NVIDIA NVML
- Virtual Environments: VMware, KVM

Motivating Problem

Hardware has become some complex that developers do not understand what it does and how it does it.

As a result it is not obvious:

- What to optimize
- How to optimize it
- How to measure the problem

Sometimes it's hard to tell what all this information means.

Is it really that hard?

How many instructions does my program execute?

INSTRUCTIONS_EXECUTED

Is it really that hard?

How many instructions does my program execute?

~~INSTRUCTIONS_EXECUTED~~

INSTRUCTION_RETIRED

BR_INST_EXEC

BR_INST_RETIRED

How many L2 misses does my kernel cause?

LLC_REFERENCES - L2_RQSTS:CODE_RD_MISS

Is it really that hard?

How many instructions does my program execute?

~~INSTRUCTIONS_EXECUTED~~

INSTRUCTION_RETIRED

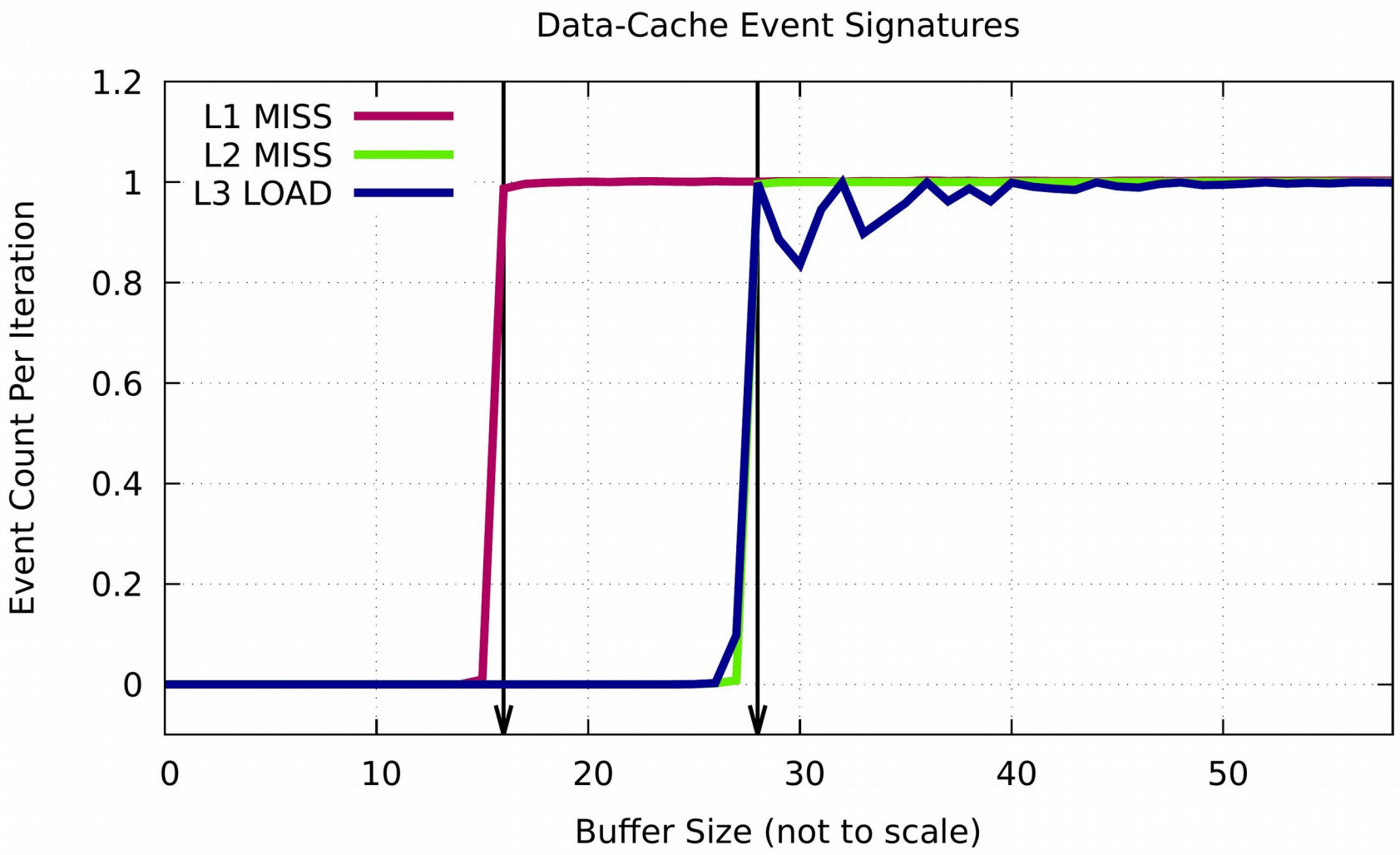
BR_INST_EXEC ← Includes speculative execution

BR_INST_RETIRED ← Executed and committed to state

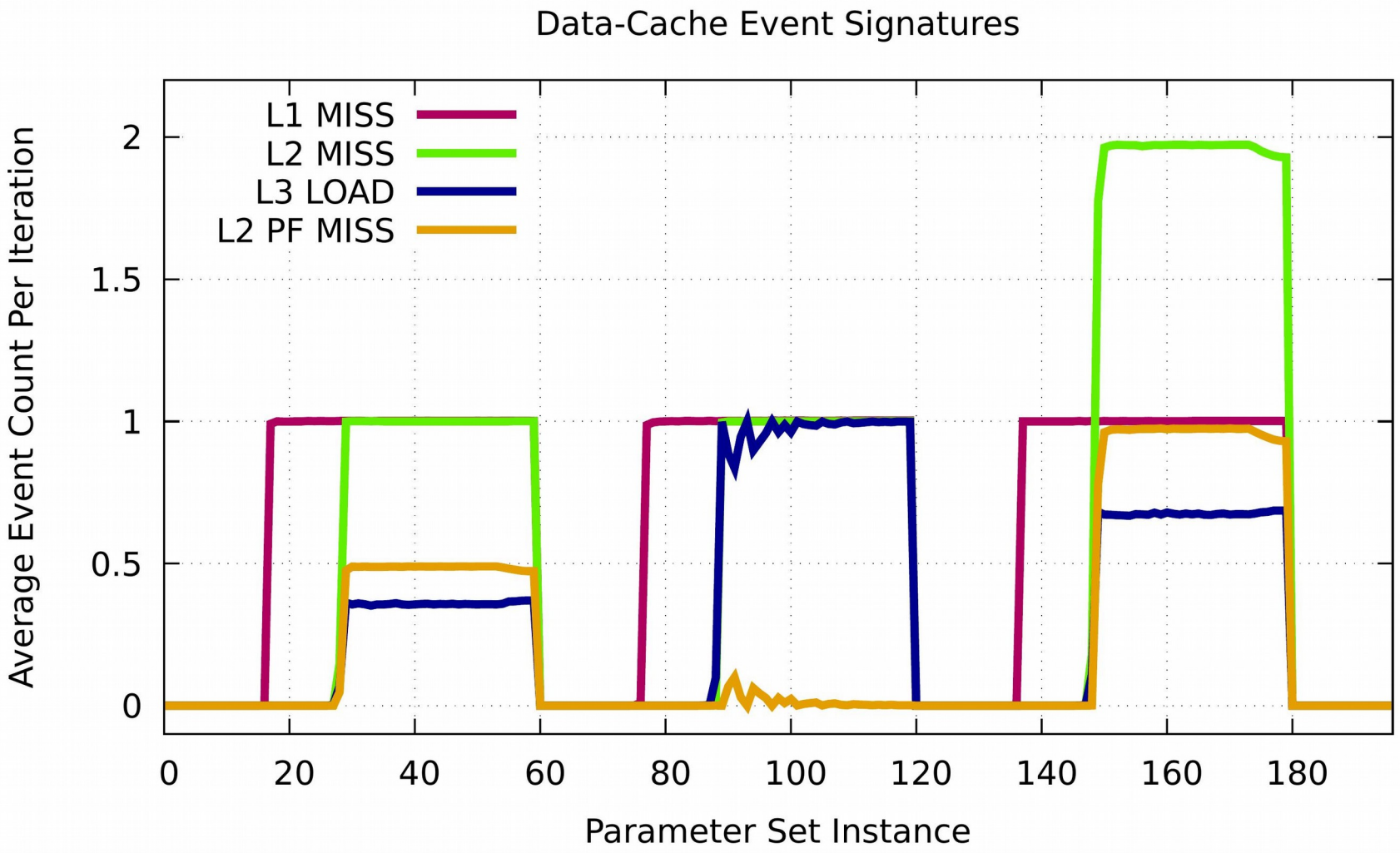
How many L2 misses does my kernel cause?

LLC_REFERENCES - L2_RQSTS:CODE_RD_MISS

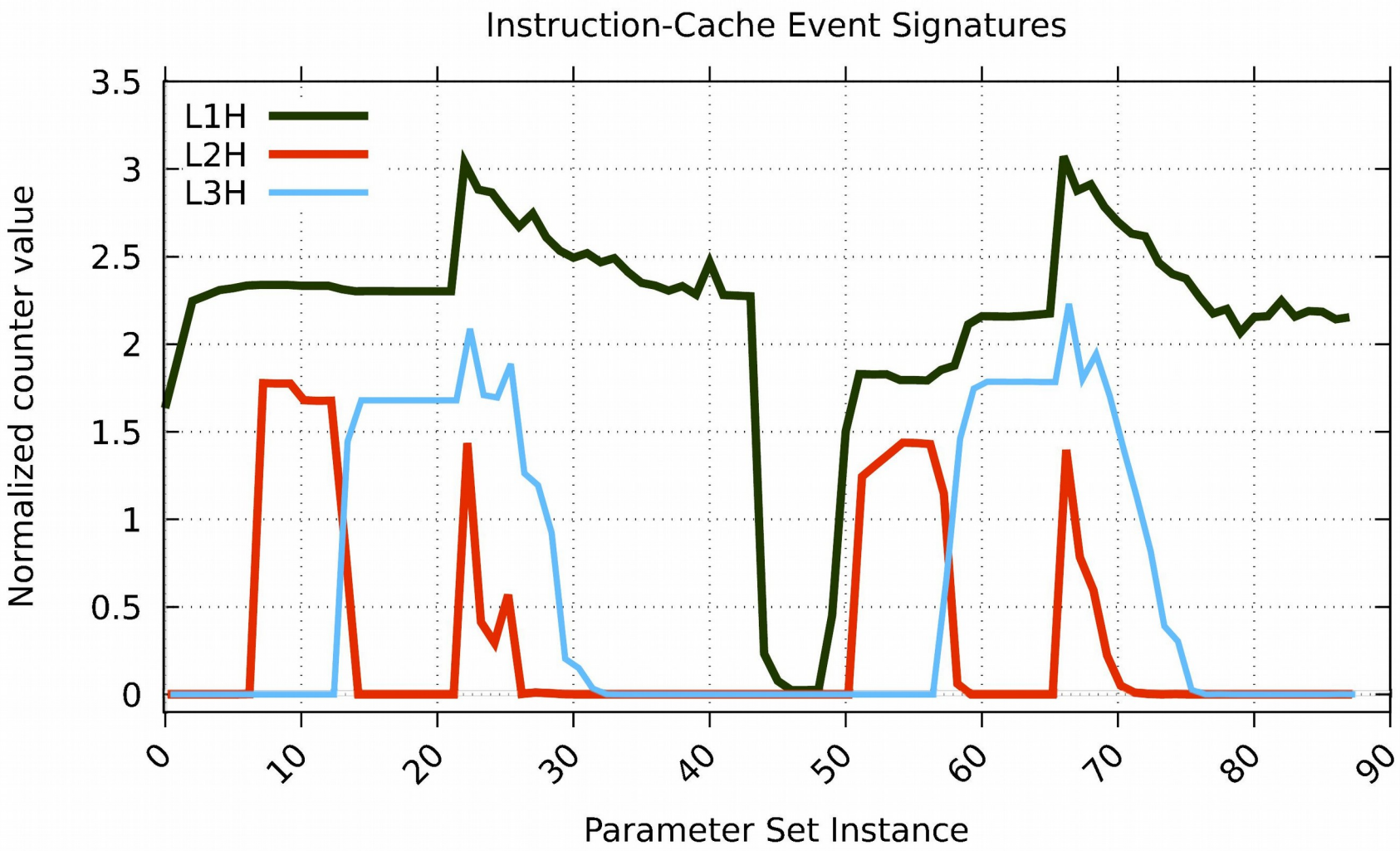
Generating event signatures



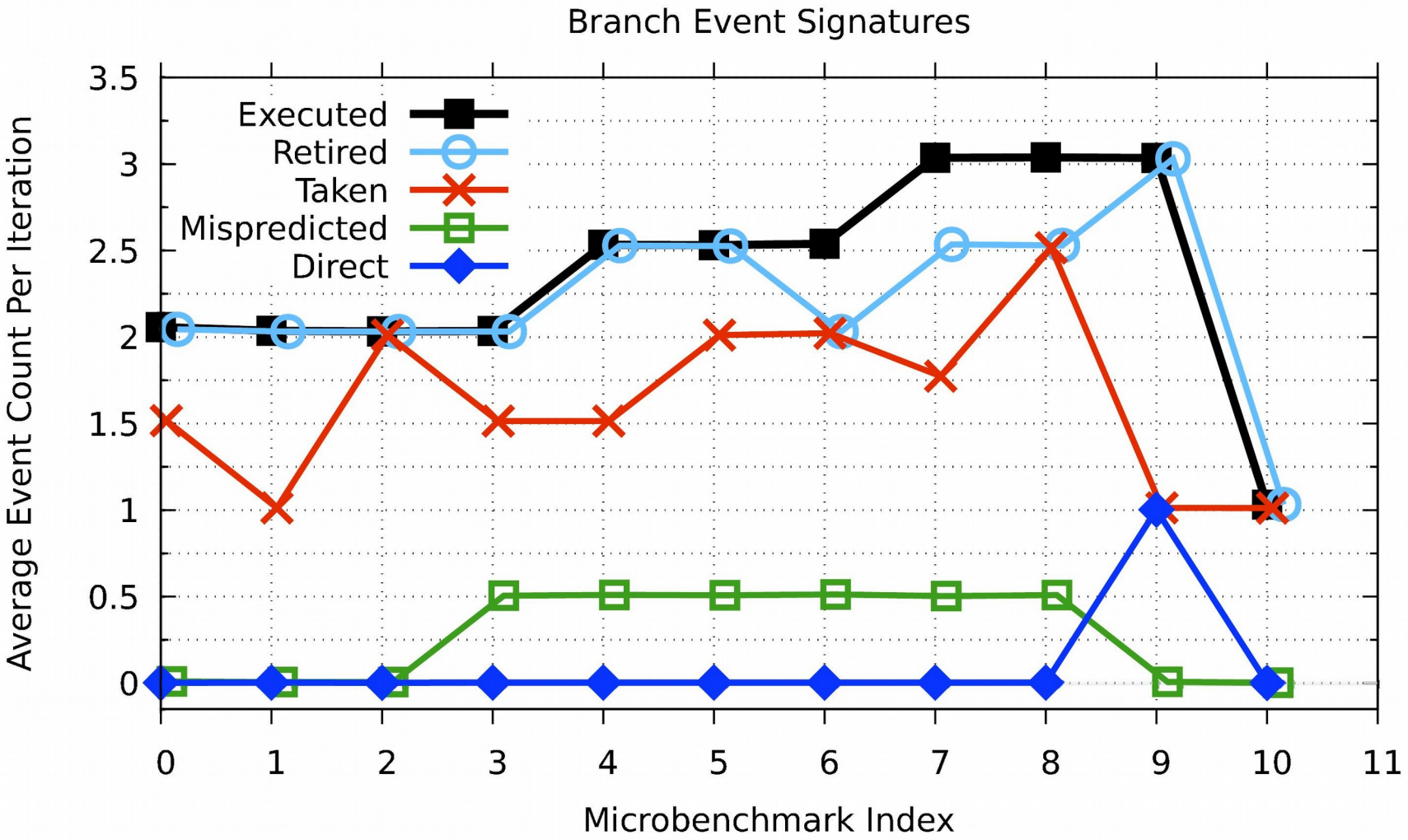
Data-cache event signatures



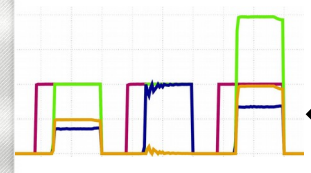
Instruction-cache event signatures

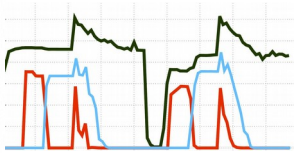


Branch event signatures



Data Analytics





$$\frac{1}{N}\sum_i \frac{(P_i - M_i)^2}{\overline{P}\cdot\overline{M}}, \quad \overline{P} = \frac{1}{N}\sum_i P_i, \quad \overline{M} = \frac{1}{N}\sum_i M_i$$

$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-\frac{i2\pi}{N}kn} \rightarrow \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} {A_i}^2}\sqrt{\sum_{i=1}^{n} {B_i}^2}}$$

L2H
L2_RQSTS:DEMAND_DATA_RD_HIT

L2M
ix86arch::LLC_REFERENCES
L2_LINES_OUT:DEMAND_CLEAN
L2_RQSTS:ALL_DEMAND_MISS
L2_RQSTS:DEMAND_DATA_RD_MISS
LONGEST_LAT_CACHE:REFERENCE
MEM_LOAD_UOPS_RETIRED:L2_MISS
perf::LLC-LOADS
OFFCORE_RESPONSE_0:DMND_DATA_RD:ANY_RESPONSE

L3H
MEM_LOAD_UOPS_L3_HIT_RETIRED:XSNP_NONE
MEM_LOAD_UOPS_LLC_HIT_RETIRED:XSNP_NONE
MEM_LOAD_UOPS_RETIRED:L3_HIT
OFFCORE_RESPONSE_0:DMND_DATA_RD:L3_HIT:SNP_ANY

L3M
ix86arch::LLC_MISSES

Summary

- Hardware is complex
- … and gets more complex with every generation
- CIT: PAPI effort that aims to alleviate complexity by
  - Categorizing, Validating, and Understanding native events
- CIT deploys micro-benchmarks and data analytics