Top-down Analysis

Working Group Outbrief

Scalable Tools Workshop 2025

Top-Down Status in current Profiling Tools

- Caliper:
  - Haswell Level 1-3
  - Sapphire Rapids: Currently requires PAPI build w/o rdpmc. Will switch to PAPI 7.2
- HPCToolkit: Level 1-4 (sampling), Sapphire Rapids (or newer)
- Likwid: Level 1 or 1-2 for Sapphire Rapids
- Linux perf stat: All levels, default level 1-2
- PAPI 7.2: Level 1-2, converts raw metrics to user-friendly percentages
- Score-P: Level 1-2
- TAU
- VTune: Level 1-4 & some level 5-6

Group’s Wishlist

- Top-Down support on other architectures such as GPU, …
- GPU:
  - Distinguish **active**, **stalled** and **idle** threads across all levels of execution units: device, blocks, warp levels
  - GPU Stalls
    - Warp threads idle due to branching
    - Memory hierarchy stalls a la Intel TMA Level 2+ metrics
  - NVidia GPUs: NC State DrGPU paper: similar to IBM Power CPI stack
  - Want to look both at whole device occupancy / efficiency as well as root-cause analysis for individual kernels
- Top-down metrics for multiple sockets, memory hierarchy
  - Uncore events: potential security issues

Usage of Top-Down Analysis

- Good overview for general behavior of the code / job-specific system monitoring.
  - More detail than rooflines.
- Give top-down data to LLM to figure out issues? https://arxiv.org/pdf/2503.13772
- LLNL uses top-down metrics to categorize application kernels into different clusters
- Top-down analysis to help for performance prediction (Probir)

Presentation of Top-down Metrics

- HPCToolkit: Use doughnut graph
  - Easy to view whole top-down hierarchy but hard to compare different program regions / call paths
- Caliper: Shows percentages. Works for 1 top-down level, multiple levels are hard to understand
- Cons of top-down analysis:
  - Generally may be difficult to understand for end users.
  - User's mental model is often oversimplified.
  - Still difficult to derive actionable improvements out of the information.
  - Involve vendors (Intel) to get directions for improving code. Are improvements portable between architectures/vendors?
- HPCToolkit and Caliper distinguish by code region but regions need to be running long enough (O(ms))

Action Plan

- Export data to JSON/TXT for ingestion in 3rd party analysis tools, LLM, …
- Need a way to quantify accuracy/uncertainty
- Talk to AMD about missing HW features to be able to do Top-Down analysis on GPUs

Resources

- Hardware support
  - Intel Top-down
    - Intel perfmon JSON
  - ARM Top-down analysis
    - NVIDIA Grace
  - IBM CPI stack : Power9 PMU
  - NVIDIA
  - AMD
- Test cases
  - Profiling games applications with top-down analysis

Participants

David Boehme

David Montoya

Probir Roy

Edgar Leon

Thomas Gruber

Laksono Adhianto

Marty Itzkowitz

Kathleen Shoga

Terry Jones