Jan 11, 2024 | RISC-V Perf Analysis SIG Meeting

Attendees: Beeman Strong tech.meetings@riscv.org Marc Casas

Notes

- **Attendees**: MattT, Guillem, RobertC, Snehasish, PaulC, DmitryR, ChunL, Keeran, DmitryP, AtishP
- **Slides/video** here
- Intros:
  - Paul Clarke - Ventana, SW perf team
  - Keeran - Rivos, SW perf team, background in profiling
  - Dmitriy Prokhorov - Syntacore, DmitryR's manager, perf analysis technologies
  - Guillem - Gaisler
  - RobertC - at MIPS now, looking at debug, trace, perfmon
- Updates
  - CTR is ARC approved pending 2 small updates
- PMU usage in profilers - DmitryR
  - Vtune background
  - Analysis starts with:
    - Hotspots - precise sampling helps
    - Call-stacks
  - Even 1 instr of skid can be misleading
  - PEBS imprecision w.r.t. Sample rate is not important, more important to get the right instr associated. Differences in event numbers across insts is often large.
  - Is there any downside to using PEBS?
    - Not really
    - Doesn't always work in a guest, until recent implementations.
      - For RISC-V we'll want to plan for virt from the beginning
  - Vtune uses LBR as auxiliary call-stack, to extend Linux perf dwarf stack
    - LBR call-stack too unreliable to use by itself, error prone, e.g. due to SW exceptions
    - But helps with some cases where large array allocated on stack
    - (Missed deeper question about dwarf stack details…)
    - By default Vtune never uses only LBR call-stack, so CTR RASEMU not likely to avoid the need to invoke SW
    - Shadow stack avoids reliability issues thanks to instrumentation
  - Problems with insufficient event documentation
    - How/when to use an event?
    - Prefer metrics, ideally standardized (e.g., TMA)
    - Goal for Performance Events TG, getting started soon
  - Memory is the most common bottleneck
    - Intel's PEBS DLA captured virtual address associated with sample, helpful

        - LLC miss, loads, & stores typically use sample rate of ~100,000
        - DLA, source, latency are easy to get with instr sampling, hard with event sampling
        - For sharing issues, Vtune doesn't use PEBS DATASRC, but does use events that indicate hit of dirty lines
    - SoC monitoring
        - Intel's OFFCORE_RESPONSE event counts SoC-level events in the core, have to dedicate wires on bus to pass back event info
    - For HPC, counting FLOPS is necessary, need dedicated events
        - AI similar, but more than FP
        - SoC monitoring important to monitor accelerators
        - FLOPS are hard to map to HW counters, moves and masks are challenging
    - What's the target audience?
        - SW devs, FW devs, HW design event
        - Less for datacenter management
    - What are use cases for FW using counters?
        - In Intel used counters to select SW class
        - Can use counters to profile FW just like any other SW
    - Some low-level debuggers use counters
    - What about system-wide profiling
        - Vtune has special platform profiler tool
        - Uses counting instead of sampling, get "over time" performance picture
        - AWS users found that some system-level counters were in use, and counts were broken sometimes in virtual machines
    - For datacenter, profiling applications relies on having environment set up, which adds challenges
        - Workloads depend on services, may be cross-node
        - Will discuss more in Snehasish's talk next week
    -

Action items

- ☐ Beeman Strong - Jul 28, 2022 - Reach out about proprietary performance analysis tools
- ☐ Beeman Strong - Jul 28, 2022 - Reach out to VMware about PMU enabling