Dec 14, 2023 | RISC-V Perf Analysis SIG Meeting

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

Notes

- **Attendees**: Beeman, BruceA, MattT, DmitryR, Snehasish, ChunL, JohnS, GregF, Ved
- **Slides/video** here
- Updates:
  - Have upcoming talks from Dmitry (analysis tool perspective) and Snehasish (Google usages). Looking for more, to help inform our priority usages. Reach out to Beeman if you have ideas.
- Resuming Sampling discussion with event sampling
  - Akin to interrupt-based sampling, but with HW assistance to make it precise, and to speed up sample state collection
  - Costly to implement (bits in ROB, staging of state, timing challenges, …)
  - Dmitry: PEBS is useful for non-precise events too
    - Then not as costly
  - There are tricks to make precise sampling cheaper
    - E.g., counter overflow arms the mechanism, collect sample on the next(ish) event
- Sampling usages
  - Began a list of priority usages to inform our decision-making
    - Hot insts, hot paths/branches, rare events, ratios, …
  - False sharing detection is one to add
  - Snehasish: for control transfers, BOLT in particular, use both CTR and CS
    - Also, Google does periodic collection of counters across datacenter, see GWP paper from 2011, and ISCA 2015, and FB from 2022. CS is essential to aggregate info across applications, for common libraries.
    - CS is a big challenge, will need to talk about how best to get that. With PEBS requires taking an interrupt on each sample.
      - ARM CSRE memcpy's call-stack? Seems like they dropped that.
  - Bruce: consider multi-core interactions. KUtrace covers these with lightweight trace.
  - Dmitry: data source, useful for false sharing if includes coherence info
    - For false sharing, sample on remote HITM accesses
    - HITM important for top-down too, contested accesses
    - Data source much easier to get with inst sampling than event sampling
  - Hoping to build this list over time
- Output buffer
  - Presume will store state to memory
  - Virtual, physical, and/or guest physical addressing?
  - Virtual good for monitoring 1 context, hard for hypervisor monitoring guest(s)
  - For monitoring across guests, need physical addressing

        - Would use guest physical for guest monitoring itself
        - Ved: SW may want to keep buffers per-context, like they do for LBRs in system-wide collections
    - Ved: could malloc the buffer address as normal, then get the PA and use it
    - Greg: SW would alloc a buffer for instr sampling, how different is this?
        - Would be good to see all the complications and interactions listed
        - Beeman: will review this in more depth in a future meeting
    - Ved: consider buffer mechanism for this and for self-hosted trace, hopefully a common mechanism
        - Greg: ditto
- Bruce: DenisB (easyperf.net) could be an invite
- Greg: scheduling, could use priv ISA slots
    - Beeman: looked at Apps & Tools slots, will consider options

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