Oct 31, 2024 | RV Performance Event Sampling TG

Attendees: Snehasish Kumar tech.meetings@riscv.org Beeman Strong

Notes

- **Attendees**: Beeman, Snehasish, Bruce, Robert, Daniel, Chun, Bo, Atish, Frank
- **Slides/video** here
- Propose that counter freeze work like shpms*: if CDE=1, only delegated counters cause freeze or get frozen
    - Prevents S-mode from freezing M-mode counters
- CTR is more complicated, LCOFIFRZ has no dependence on CDE
    - If any counter overflows/LCOFIs when LCOFIFRZ=1, CTR will freeze
    - If there are 2 PMU users (e.g., hypervisor and guest), can get some undesirable freeze behavior. Discussed examples.
    - Ideally CTR freeze would be counter-specific
- Propose that, with Sshpmfrz (name for counter freeze extension), only counters selected in scounterfrz cause CTR to freeze
    - Means CTR shares freeze list with counter freeze, though each can be enabled separately. I.e., can use list to freeze only counters, or only CTR, or both
- Does Sshpmfrz become required to make CTR freeze usable?
    - Arguably. Though x86 LBR freeze is global, guess it's handled by perf somehow?
    - Counter freeze is not supported by perf for x86, Atish will check on LBR freeze
- Are CTR freeze examples, with hypervisor and guest each using counters, assuming pass-through PMU?
    - No, could be emulated (SBI) ifc
- How do we clear freeze? Are there race conditions
    - Forgot to include this in the slides
    - Propose that if any OF bit for a selected counter is set, all selected counters are inhibited. Once handler clears OF bits for all selected counters, counters resume counting.
        - So no HW updates to scounterfrz
        - Don't foresee any race conditions, since scounterfrz and OF bits should be written only at PMU init
        - Perf handler clears OF bits only when re-starting the counter, so this should work
    - An alternative would be for freeze to update xcountinhibit, setting the bits for selected counters to inhibit them. Then handler clears those bits (as it does already).
        - That could produce more races, with SW writing xcountinhibit
- Will need to save/restore scounterfrz on PMU switch
- Some implementations could have freeze skid, where it takes a 1-2 cycles to freeze. Should be less than LCOFI skid.

- Should we bother with freeze?
    - Value is modest
        - Useful to avoid counting in the handler for kernel profiling
        - Useful to simplify identifying precise PC for architectural events (like INST.RET)
    - If perf doesn't support it for x86 maybe ecosystem isn't really there for it?
    - Does ARM have freeze? Beeman to look
- Discussed issues with counter sharing
    - Only 1 freeze group, sharing users are stuck with the same freeze timing
    - Could consider adding a vscounterfrz, to allow hypervisor and guest to have independent freeze groups/configs
- GWP does concurrent sampling
    - Have a primary event, and cycles/instructions. Cycles isn't used, instret is, for PKI metrics per PC
    - Leader sampling only used individually, not fleet-wide. Easy to get wrong. Something like cycles may be biased to longer latency insts.
    - Sampling on multiple counters reduces load on infra, use the sample data independently
- So probably unwise to support sampling or precise attribution on just 1 counter
- Event list should make it straightforward to indicate which counters support non-spec events and precise attribution/LCOFI
- Next time will review freeze clearing proposal, and freeze value. Then can start on instr sampling.

Action items

- ☐ Atish Kumar Patra - 10/31/2024 - check on perf support for LBR freeze
- ☐ Beeman Strong - 10/31/2024 - check on ARM support for freeze