



Linux Plumbers Conference

Richmond, Virginia | November 13-15, 2023



Linux
Plumbers
Conference | Richmond, VA | Nov. 13-15, 2023

Perf improvements in RISC-V

Atish Patra <Rivos>





Current Status

- SBI PMU extension + Sscofpmf enables perf stat/record usage
- Some existing hardware platform (Andes/T-head) has similar implementation as Sscofpmf but different CSR/interrupt number
- These platforms work with alternate framework
- Event discovery happens through a firmware only DT node



Upcoming changes – ISA extensions

- Upcoming ISA feature extensions (Smcdeleg, Ssccfg) enables access/program these counters in S-mode directly
- Patches [available](#)
- Counter delegation works in backward compatible manner
 - The driver uses the ISA extensions for hardware events and SBI PMU for firmware events
 - Challenges
 - Event encoding
 - Event to counter mapping



Challenges

- No standard architectural events apart from cycle/instructions
 - These events do not support sampling
 - Vendor specific programmable hpmcounter is the only solution
- Other perf standard event encoding will differ from one vendor to another
- ISA doesn't say anything about event to counter mapping



Event encoding discovery

- Any platform(including Qemu) using the Smcdeleg/Ssccfg **must** use the json file to encode the events
- Only support raw events for now
- No predefined events via sysfs
- Opens:
 - Perf standard hardware/cache events (i.e. cycles/instructions) renamed to be used as raw events
 - Need a way to map the perf standard events to map to the same events specified in the json file

```
{
  "EventName": "cycles_executed",
  "EventCode": "0x00",
  "BriefDescription": "cycle executed",
  "CounterIDMask": "0xF8"
},
{
  "EventName": "inst_retired",
  "EventCode": "0x01",
  "BriefDescription": "instruction retired",
  "CounterIDMask": "0xF8"
},
{
  "EventName": "dTLB_load_misses",
  "EventCode": "0x10019",
  "BriefDescription": "Data TLB load miss",
  "CounterIDMask": "0xF8"
},
{
  "EventName": "dTLB_store_misses",
  "EventCode": "0x1001B",
  "BriefDescription": "Data TLB store miss",
  "CounterIDMask": "0xF8"
},
{
  "EventName": "iTLB_load_misses",
  "EventCode": "0x10021",
  "BriefDescription": "Instruction fetch TLB load miss",
  "CounterIDMask": "0xF8"
}
```




Event – Counter mapping

- The ISA doesn't mandate any specific mappings between events/counters
- Other ISA Comparison
 - ARM64 allows all-to-all mapping
 - RISC-V perf analysis SIG did not agree to this to allow flexibility
 - x86 has constraints based mapping which updates based on cpu model
 - RISC-V will have many vendors. Having vendor specific functions to specify constraints not scalable
- Current approach relies based on Json
 - Introduced a new field **CounterIdMask** -> **config2**
 - **PMU_FORMAT_ATTR(counterid_mask, "config2:0-31");**
 - x86 perfmon has a field **Counters**. Not used by perf



Event encoding discovery

- No standard event encoding in ISA
- Vendor specific encoding not allowed via sysfs yet
 - SBI PMU : config (0:47)
 - Counter delegation : config (0:56)
- No other provision for vendor specific encoding details
 - x86 standard encoding fields (i.e. umask, inv, cmask, edge)
- Do we need a vendor specific encoding scheme via sysfs ?
- Any other alternative solution ?



Freeze on overflow

- Seeking early feedback for proposal. Yet to be discussed within broader RISC-V community
- Freeze on Overflow **Interrupt**
 - The feature was removed in x86 few years back
 - Freeze only the counters that overflow. Others continue to avoid the undercount.
 - Required for rr tool ?
- Freeze on counter overflow
 - Freeze all the selected counters
 - Useful for leader sampling use case
- Any preference for any of the above approach ?