

PM5/PM6

Energieabschätzung mit Hilfe von Ereigniszählern

Mit Hilfe von Ereigniszählern lässt sich die Leistungsaufnahme eines Prozessors abschätzen [1]. In dieser Aufgabe sollen einzelne Tasks durch die von ihnen verursachte Leistungsaufnahme charakterisiert werden. Dazu muss zunächst ein Mechanismus zur Energieabschätzung implementiert werden. Die Abschätzung geschieht durch Zählen von Ereignissen und durch Gewichten der Zählerwerte mit Energiegewichten.

Die Ereigniszähler des Pentium 4 sind in der Lage, eine Vielzahl an verschiedenen Ereignissen zu zählen. Unter `/home/power/pub/p4_perfcounter` liegt ein Script, welches die Zählregister so konfiguriert, dass sie die in der Tabelle aufgeführten Ereignisse zählen.

| Ereignis | Zählregister | Gewicht [nJ] |
|--------------------|-----------------------|--------------|
| time stamp counter | MSR_IA32_TS_COUNTER | 8,14 |
| unhalted cycles | MSR_P4_BPU_COUNTER0 | 11,5 |
| µop queue writes | MSR_P4_MS_COUNTER0 | 7,71 |
| retired branches | MSR_P4_MS_COUNTER1 | 0,914 |
| mispred branches | MSR_P4_IQ_COUNTER0 | 552 |
| mem retired | MSR_P4_IQ_COUNTER1 | 2,81 |
| mob load replay | MSR_P4_BPU_COUNTER1 | 48,6 |
| ld miss 1L retired | MSR_P4_IQ_COUNTER2 | 22,1 |
| floating point | MSR_P4_FLAME_COUNTER0 | 0,697 |

Zum Charakterisieren einzelner Tasks müssen am Ende jeder Zeitscheibe und bei jedem Taskwechsel (siehe `kernel/sched.c`) die Ereigniszähler ausgelesen und der Energieverbrauch ermittelt werden. Dividiert man den Energieverbrauch durch die seit dem letzten Messen verstrichene Zeit, erhält man die durch den Task verursachte Leistungsaufnahme. Zum Speichern der Leistungsaufnahme soll die Datenstruktur `task_struct` (`include/linux/sched.h`) um ein Feld `power` erweitert werden. Dieses Feld soll vom Userspace aus über das `proc`-Dateisystem auslesbar sein.

Hinweise:

- Im Linux-Kern können keine Gleitkommaoperationen ausgeführt werden. Alle Berechnungen müssen daher als Ganzzahloperationen erfolgen.
- Die Ereigniszähler umfassen 48 Bit. Unter `/home/power/pub/p4_perfcounter/rdmsr11_wrmsr11.h` finden sich Makros zum Auslesen/Schreiben dieser Register in/aus eine(r) 64-Bit-Variable(n).

- [1] Frank Bellosa, Andreas Weissel, Martin Waitz and Simon Kellner: Event-Driven Energy Accounting for Dynamic Thermal Management. In *Proceedings of the Workshop on Compilers and Operating Systems for Low Power (COLP'03)*, September 2003.
http://www4.informatik.uni-erlangen.de/Publications/pdf/Bellosa-Weissel-Kellner-Waitz-COLP03-Thermal_Management.pdf