Full credit is given to the above companies including the OS that this PDF file was generated!

Red Hat Enterprise Linux Release 9.2 Manual Pages on 'cpupower-monitor.1' command

$ man cpupower-monitor.1

CPUPOWER-MONITOR(1) cpupower Manual CPUPOWER-MONITOR(1)

NAME

cpupower-monitor - Report processor frequency and idle statistics

SYNOPSIS

cpupower monitor -l

cpupower monitor [-c][-m <mon1>,[<mon2>,...]] [-i seconds]

cpupower monitor [-c][-m <mon1>,[<mon2>,...]] command

DESCRIPTION

cpupower-monitor reports processor topology, frequency and idle power state statistics. Either command is forked and statistics are printed upon its completion, or statistics are printed periodically.

cpupower-monitor implements independent processor sleep state and fre? quency counters. Some are retrieved from kernel statistics, some are directly reading out hardware registers. Use -l to get an overview which are supported on your system.

Options

-l

List available monitors on your system. Additional details about each monitor are shown:

? The name in quotation marks which can be passed to the -m parameter.

? The number of different counters the monitor supports in brackets.

? The amount of time in seconds the counters might overflow, due to implementation constraints.

? The name and a description of each counter and its proces? sor hierarchy level coverage in square brackets:

? [T] -> Thread

? [C] -> Core

? [P] -> Processor Package (Socket)

? [M] -> Machine/Platform wide counter

-m <mon1>,<mon2>,...

Only display specific monitors. Use the monitor string(s) provided by -l option.

-i seconds

Measure interval.

-c

Schedule the process on every core before starting and ending mea? suring. This could be needed for the Idle_Stats monitor when no other MSR based monitor (has to be run on the core that is mea? sured) is run in parallel. This is to wake up the processors from deeper sleep states and let the kernel re -account its cpuidle (C-state) information before reading the cpuidle timings from sysfs.

command

Measure idle and frequency characteristics of an arbitrary com? mand/workload. The executable command is forked and upon its exit, statistics gathered since it was forked are displayed.

-v

Increase verbosity if the binary was compiled with the DEBUG option set.

MONITOR DESCRIPTIONS

Idle_Stats

Shows statistics of the cpuidle kernel subsystem. Values are retrieved from /sys/devices/system/cpu/cpu*/cpuidle/state*/. The kernel updates these values every time an idle state is entered or left. Therefore there can be some inaccuracy when cores are in an idle state for some

time when the measure starts or ends. In worst case it can happen that one core stayed in an idle state for the whole measure time and the idle state usage time as exported by the kernel did not get updated. In this case a state residency of 0 percent is shown while it was 100.

Mperf

The name comes from the aperf/mperf (average and maximum) MSR registers used which are available on recent X86 processors. It shows the average frequency (including boost frequencies). The fact that on all recent hardware the mperf timer stops ticking in any idle state it is also used to show C0 (processor is active) and Cx (processor is in any sleep state) times. These counters do not have the inaccuracy restrictions the "Idle_Stats" counters may show. May work poorly on Linux-2.6.20 through 2.6.29, as the acpi-cpufreq kernel frequency driver periodi? cally cleared aperf/mperf registers in those kernels.

Nehalem SandyBridge HaswellExtended

Intel Core and Package sleep state counters. Threads (hyperthreaded cores) may not be able to enter deeper core states if its sibling is utilized. Deepest package sleep states may in reality show up as ma? chine/platform wide sleep states and can only be entered if all cores are idle. Look up Intel manuals (some are provided in the References section) for further details. The monitors are named after the CPU family where the sleep state capabilities got introduced and may not match exactly the CPU name of the platform. For example an IvyBridge processor has sleep state capabilities which got introduced in Nehalem and SandyBridge processor families. Thus on an IvyBridge processor one will get Nehalem and SandyBridge sleep state monitors. HaswellExtended extra package sleep state capabilities are available only in a specific Haswell (family 0x45) and probably also other future processors.

Fam_12h Fam_14h

AMD laptop and desktop processor (family 12h and 14h) sleep state coun? ters. The registers are accessed via PCI and therefore can still be read out while cores have been offlined.

There is one special counter: NBP1 (North Bridge P1). This one always

returns 0 or 1, depending on whether the North Bridge P1 power state got entered at least once during measure time. Being able to enter NBP1 state also depends on graphics power management. Therefore this counter can be used to verify whether the graphics' driver power man? agement is working as expected.

EXAMPLES

cpupower monitor -l" may show:

Monitor "Mperf" (3 states) - Might overflow after 922000000 s

...

Monitor "Idle_Stats" (3 states) - Might overflow after 4294967295 s

...

cpupower monitor -m "Idle_Stats,Mperf" scp /tmp/test /nfs/tmp

Monitor the scp command, show both Mperf and Idle_Stats states counter statistics, but in exchanged order.

Be careful that the typical command to fully utilize one CPU by doing:

cpupower monitor cat /dev/zero >/dev/null

Does not work as expected, because the measured output is redirected to /dev/null. This could get workarounded by putting the line into an own, tiny shell script. Hit CTRL-c to terminate the command and get the mea? sure output displayed.

REFERENCES

"BIOS and Kernel Developer?s Guide (BKDG) for AMD Family 14h Proces? sors" https://support.amd.com/us/Processor_TechDocs/43170.pdf

"Intel? Turbo Boost Technology in Intel? Core? Microarchitecture (Ne? halem) Based Processors" http://download.intel.com/design/processor/ap? plnots/320354.pdf

"Intel? 64 and IA-32 Architectures Software Developer's Manual Volume 3B: System Programming Guide" https://www.intel.com/products/proces? sor/manuals

FILES

/dev/cpu/*/msr

/sys/devices/system/cpu/cpu*/cpuidle/state*/.

SEE ALSO

powertop(8), msr(4), vmstat(8)

AUTHORS

Written by Thomas Renninger <trenn@suse.de>

Nehalem, SandyBridge monitors and command passing

based on turbostat.8 from Len Brown <len.brown@intel.com>

22/02/2011 CPUPOWER-MONITOR(1)