Rocky Enterprise Linux 9.2 Manual Pages on command 'x86_energy_perf_policy.8'

$ man x86_energy_perf_policy.8

X86_ENERGY_PERF_POLICY(8) System Manager's Manual X86_ENERGY_PERF_POLICY(8)

NAME

x86_energy_perf_policy - Manage Energy vs. Performance Policy via x86

Model Specific Registers

SYNOPSIS

x86_energy_perf_policy [options] [scope] [field value]

scope: --cpu cpu-list | --pkg pkg-list

cpu-list, pkg-list: # | #,# | #-# | all

field: --all | --epb | --hwp-epp | --hwp-min | --hwp-max | --hwp-de?

sired

other: (--force | --hwp-enable | --turbo-enable) value)

value: # | default | performance | balance-performance | balance-power

| power

DESCRIPTION

x86_energy_perf_policy displays and updates energy-performance policy

settings specific to Intel Architecture Processors. Settings are ac?

cessed via Model Specific Register (MSR) updates, no matter if the

Linux cpufreq sub-system is enabled or not.

Policy in MSR_IA32_ENERGY_PERF_BIAS (EPB) may affect a wide range of hardware decisions, such as how aggressively the hardware enters and exits CPU idle states (C-states) and Processor Performance States (P-states). This policy hint does not replace explicit OS C-state and P-state selection. Rather, it tells the hardware how aggressively to im? plement those selections. Further, it allows the OS to influence en? ergy/performance trade-offs where there is no software interface, such as in the opportunistic "turbo-mode" P-state range. Note that MSR_IA32_ENERGY_PERF_BIAS is defined per CPU, but some implementations share a single MSR among all CPUs in each processor package. On those systems, a write to EPB on one processor will be visible, and will have an effect, on all CPUs in the same processor package.

Hardware P-States (HWP) are effectively an expansion of hardware P-state control from the opportunistic turbo-mode P-state range to in? clude the entire range of available P-states. On Broadwell Xeon, the initial HWP implementation, EPB influenced HWP. That influence was re? moved in subsequent generations, where it was moved to the Energy_Per? formance_Preference (EPP) field in a pair of dedicated MSRs -- MSR_IA32_HWP_REQUEST and MSR_IA32_HWP_REQUEST_PKG.

EPP is the most commonly managed knob in HWP mode, but MSR_IA32_HWP_RE? QUEST also allows the user to specify minimum-frequency for Quality-of-Service, and maximum-frequency for power-capping. MSR_IA32_HWP_REQUEST is defined per-CPU.

MSR_IA32_HWP_REQUEST_PKG has the same capability as MSR_IA32_HWP_RE? QUEST, but it can simultaneously set the default policy for all CPUs within a package. A bit in per-CPU MSR_IA32_HWP_REQUEST indicates whether it is over-ruled-by or exempt-from MSR_IA32_HWP_REQUEST_PKG.

MSR_HWP_CAPABILITIES shows the default values for the fields in MSR_IA32_HWP_REQUEST. It is displayed when no values are being writ? ten.

SCOPE OPTIONS

-c, --cpu Operate on the MSR_IA32_HWP_REQUEST for each CPU in a CPU-list. The CPU-list may be comma-separated CPU numbers, with dash for

range or the string "all". Eg. '--cpu 1,4,6-8' or '--cpu all'. When

--cpu is used, --hwp-use-pkg is available, which specifies whether the

per-cpu MSR_IA32_HWP_REQUEST should be over-ruled by MSR_IA32_HWP_RE?

QUEST_PKG (1), or exempt from MSR_IA32_HWP_REQUEST_PKG (0).

-p, --pkg Operate on the MSR_IA32_HWP_REQUEST_PKG for each package in

the package-list. The list is a string of individual package numbers

separated by commas, and or ranges of package numbers separated by a

dash, or the string "all". For example '--pkg 1,3' or '--pkg all'

VALUE OPTIONS

normal | default Set a policy with a normal balance between performance

and energy efficiency. The processor will tolerate minor performance

compromise for potentially significant energy savings. This is a rea?

sonable default for most desktops and servers. "default" is a synonym

for "normal".

performance Set a policy for maximum performance, accepting no perfor?

mance sacrifice for the benefit of energy efficiency.

balance-performance Set a policy with a high priority on performance,

but allowing some performance loss to benefit energy efficiency.

balance-power Set a policy where the performance and power are bal?

anced. This is the default.

power Set a policy where the processor can accept a measurable perfor?

mance impact to maximize energy efficiency.

The following table shows the mapping from the value strings above to

actual MSR values. This mapping is defined in the Linux-kernel header,

msr-index.h.

VALUE STRING EPB EPP

performance 0 0

balance-performance 4 128

normal, default 6 128

balance-power 8 192

power 15 255

For MSR_IA32_HWP_REQUEST performance fields (--hwp-min, --hwp-max,

--hwp-desired), the value option is in units of 100 MHz, Eg. 12 signi?

fies 1200 MHz.

FIELD OPTIONS

-a, --all value-string Sets all EPB and EPP and HWP limit fields to the value associated with the value-string. In addition, enables turbo-mode and HWP-mode, if they were previous disabled. Thus "--all normal" will set a system without cpufreq into a well known configuration.

-B, --epb set EPB per-core or per-package. See value strings in the table above.

-d, --debug debug increases verbosity. By default x86_energy_perf_pol?icy is silent for updates, and verbose for read-only mode.

-P, --hwp-epp set HWP.EPP per-core or per-package. See value strings in the table above.

-m, --hwp-min request HWP to not go below the specified core/bus ratio. The "default" is the value found in IA32_HWP_CAPABILITIES.min.

-M, --hwp-max request HWP not exceed a the specified core/bus ratio. The "default" is the value found in IA32_HWP_CAPABILITIES.max.

-D, --hwp-desired request HWP 'desired' frequency. The "normal" set?ting is 0, which corresponds to 'full autonomous' HWP control. Non-zero performance values request a specific performance level on this processor, specified in multiples of 100 MHz.

-w, --hwp-window specify integer number of microsec in the sliding win?dow that HWP uses to maintain average frequency. This parameter is meaningful only when the "desired" field above is non-zero. Default is 0, allowing the HW to choose.

OTHER OPTIONS

-f, --force writes the specified values without bounds checking.

-U, --hwp-use-pkg (0 | 1), when used in conjunction with --cpu, indi?cates whether the per-CPU MSR_IA32_HWP_REQUEST should be overruled (1) or exempt (0) from per-Package MSR_IA32_HWP_REQUEST_PKG settings. The default is exempt.

-H, --hwp-enable enable HardWare-P-state (HWP) mode. Once enabled, system RESET is required to disable HWP mode.

-t, --turbo-enable enable (1) or disable (0) turbo mode.

-v, --version print version and exit.

If no request to change policy is made, the default behavior is to read

and display the current system state, including the default capabili?

ties.

WARNING

This utility writes directly to Model Specific Registers. There is no

locking or coordination should this utility be used to modify HWP limit

fields at the same time that intel_pstate's sysfs attributes access the

same MSRs.

Note that --hwp-desired and --hwp-window are considered experimental.

Future versions of Linux reserve the right to access these fields in?

ternally -- potentially conflicting with user-space access.

EXAMPLE

```
# sudo x86_energy_perf_policy
cpu0: EPB 6
cpu0: HWP_REQ: min 6 max 35 des 0 epp 128 window 0x0 (0*10^0us) use_pkg 0
cpu0: HWP_CAP: low 1 eff 8 guar 27 high 35
cpu1: EPB 6
cpu1: HWP_REQ: min 6 max 35 des 0 epp 128 window 0x0 (0*10^0us) use_pkg 0
cpu1: HWP_CAP: low 1 eff 8 guar 27 high 35
cpu2: EPB 6
cpu2: HWP_REQ: min 6 max 35 des 0 epp 128 window 0x0 (0*10^0us) use_pkg 0
cpu2: HWP_CAP: low 1 eff 8 guar 27 high 35
cpu3: EPB 6
cpu3: HWP_REQ: min 6 max 35 des 0 epp 128 window 0x0 (0*10^0us) use_pkg 0
cpu3: HWP_CAP: low 1 eff 8 guar 27 high 35
```

NOTES

x86_energy_perf_policy runs only as root.

FILES

/dev/cpu/*/msr

SEE ALSO

msr(4)

Intel(R) 64 and IA-32 Architectures Software Developer's Manual

AUTHORS

Len Brown