Rock and a Hard Place

How Hard It Is To Be a CPU Idle Time Governor

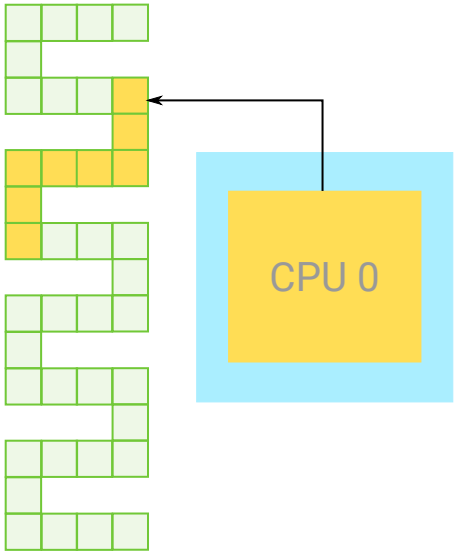
Rafael J. Wysocki

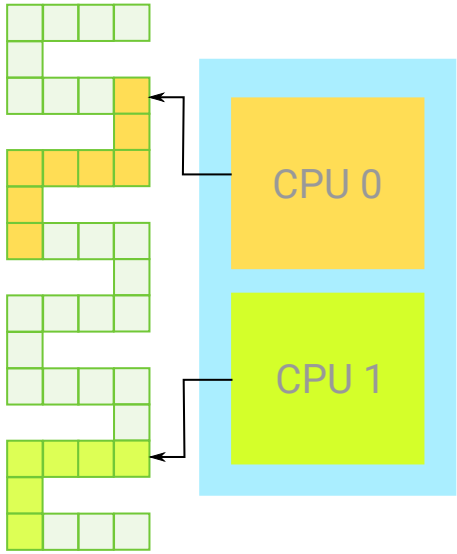
Intel Linux Systems Engineering

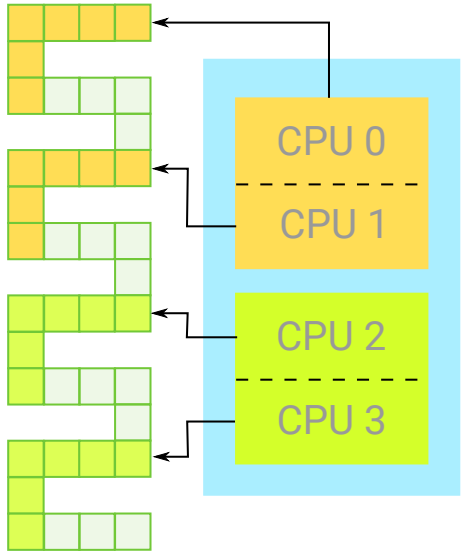
May 20, 2019

Intel OpenSource TECHNOLOGY CENTER

Terminology: CPUs = Logical CPUs







Intel OpenSource TECHNOLOGY CENTER

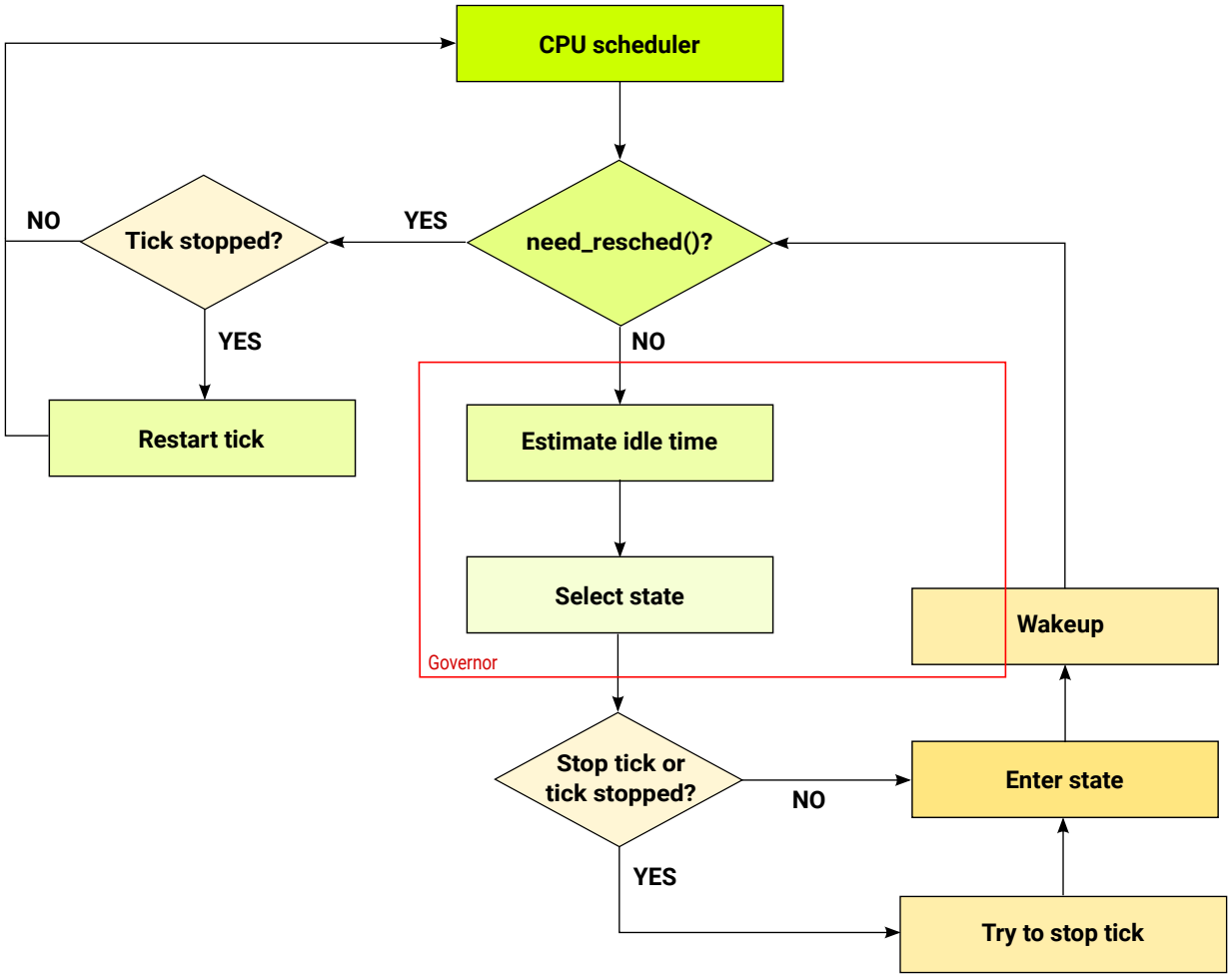
CPUs: Busy Vs Idle

Intel OpenSource TECHNOLOGY CENTER

Idleness Is an Opportunity to Save Energy

Intel OpenSource TECHNOLOGY CENTER

The Idle Loop in Linux* (4.17 and Later)



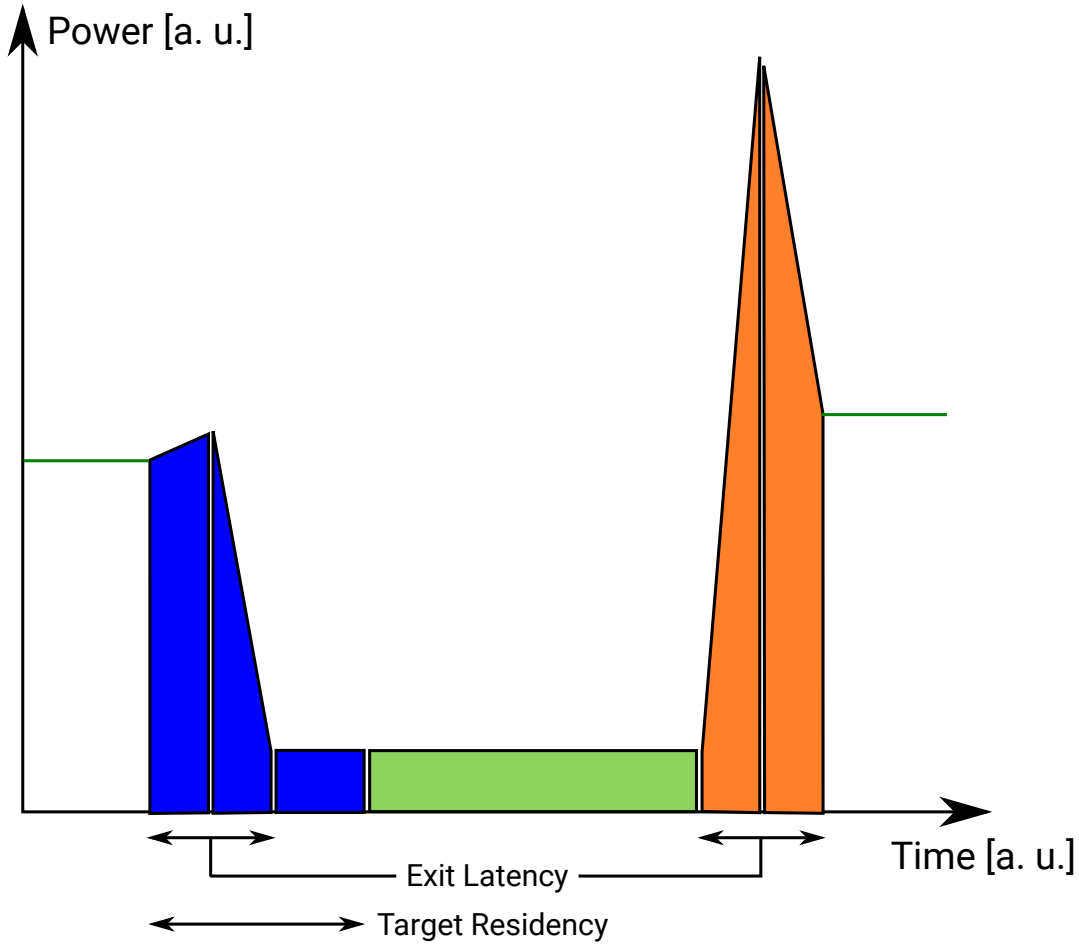
Intel OpenSource TECHNOLOGY CENTER

* Other names and brands may be claimed as the property of others

Idle Governor and Idle Driver Need To Communicate

Intel OpenSource TECHNOLOGY CENTER

CPU Idle State Parameters



Intel OpenSource TECHNOLOGY CENTER

Worst-case Numbers Are Difficult To Come By

Worst case measurement Catch 22

In order to measure the worst case you must experience it, but you cannot know whether or not you have experienced it unless you know beforehand what it is.

Intel OpenSource TECHNOLOGY CENTER

Governor Perspective

Intel OpenSource TECHNOLOGY CENTER

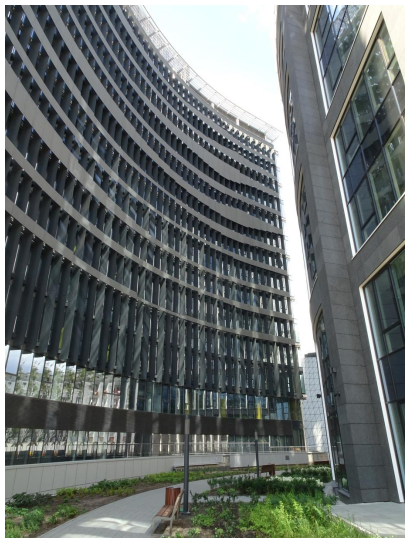
The Rock and The Hard Place

The Rock

Selecting idle states that are too shallow hurts energy-efficiency.

The Hard Place

Selecting idle states that are too deep hurts energy-efficiency and performance (through excessive latency).



Intel OpenSource TECHNOLOGY CENTER

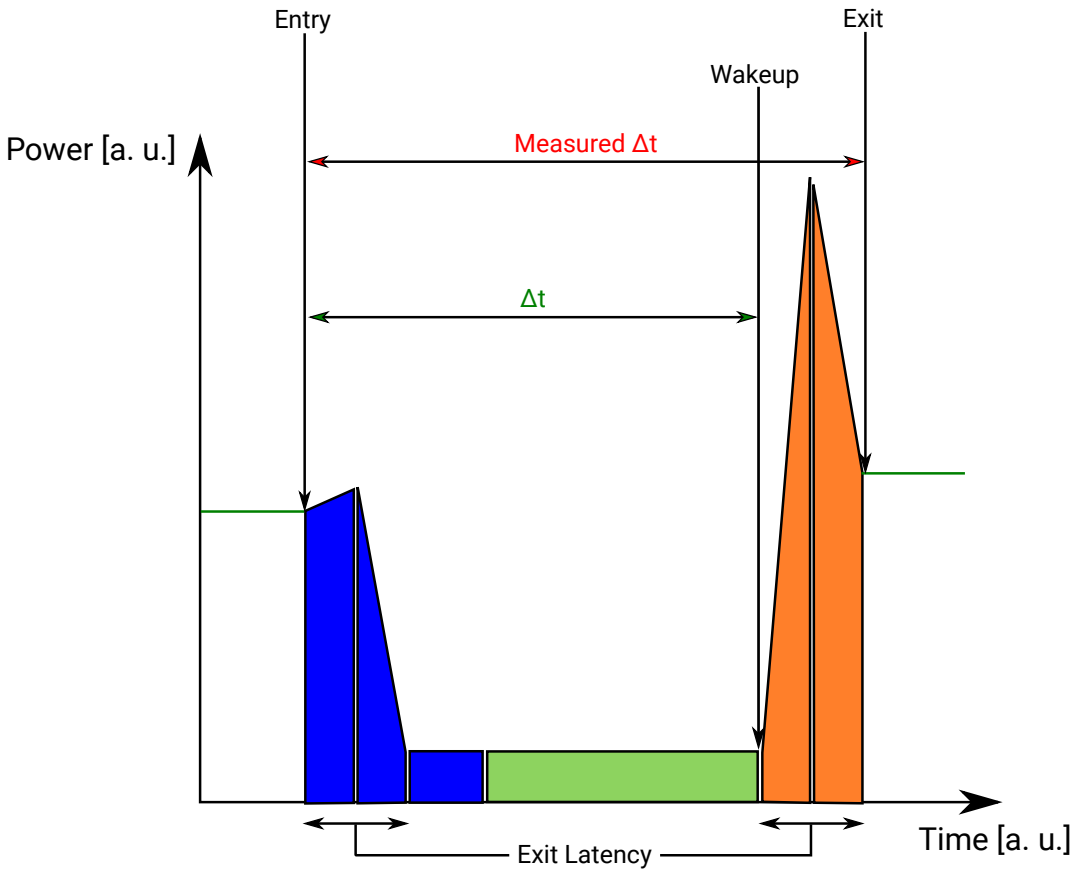
Two Categories of Wakeup Events

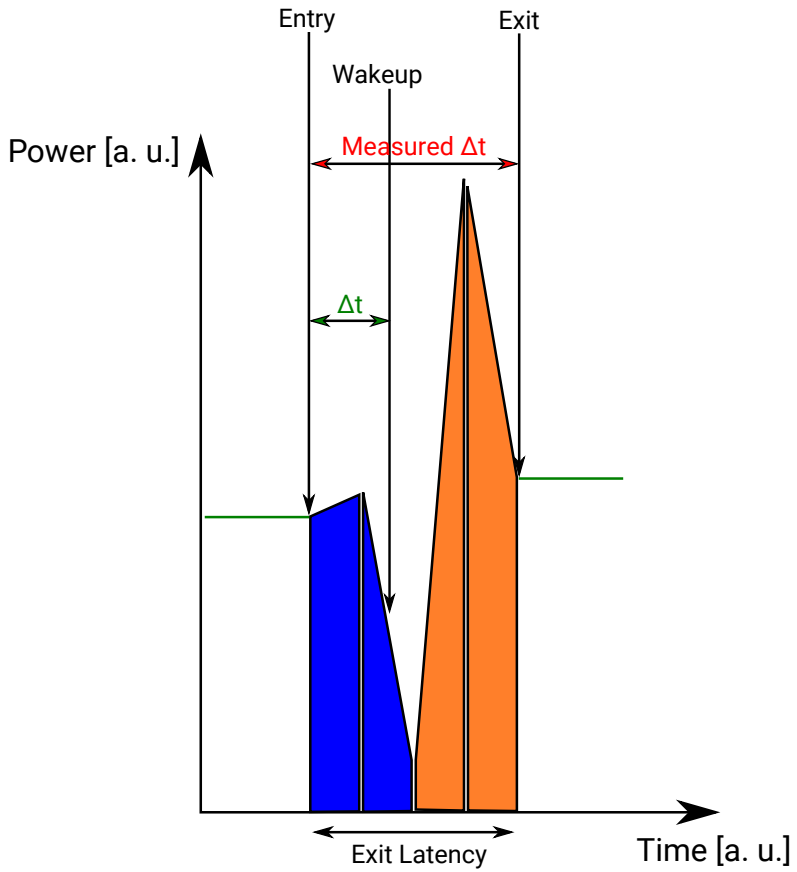
Intel OpenSource TECHNOLOGY CENTER

Measurement Issues

Intel OpenSource TECHNOLOGY CENTER

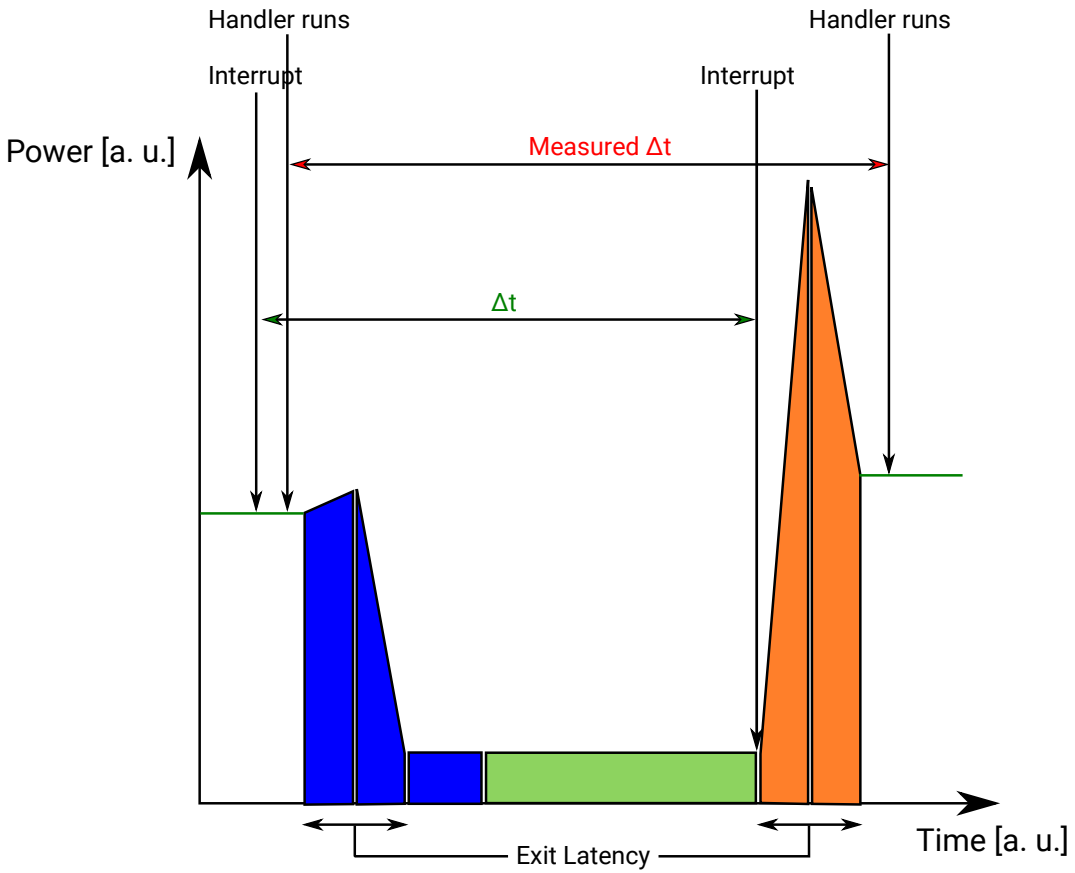
Idle Duration Is Difficult To Measure Precisely

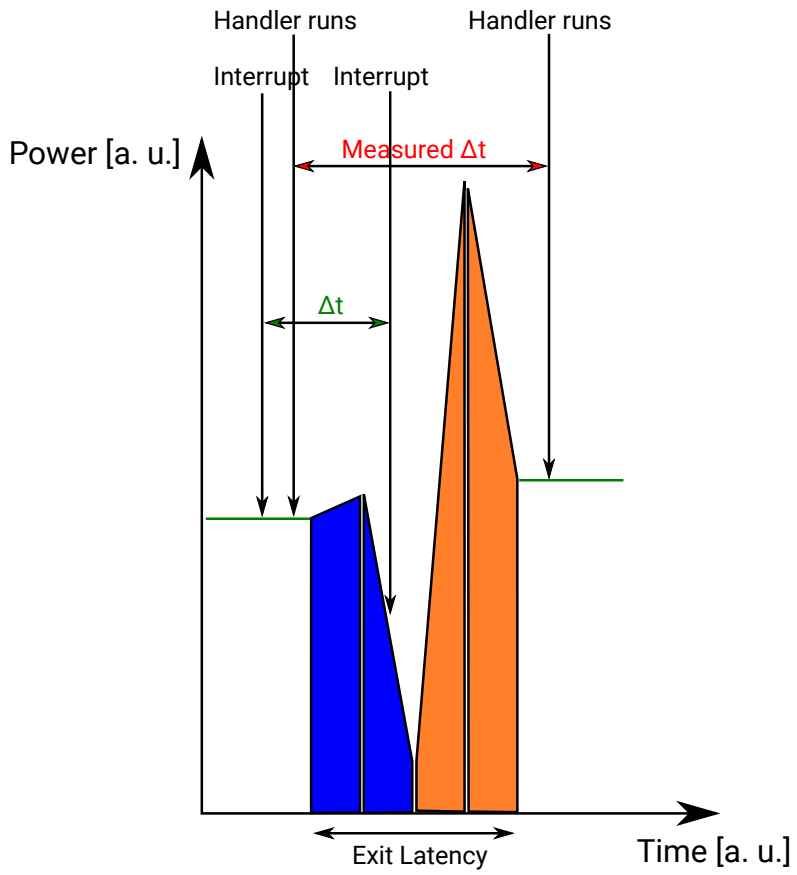




Intel OpenSource TECHNOLOGY CENTER

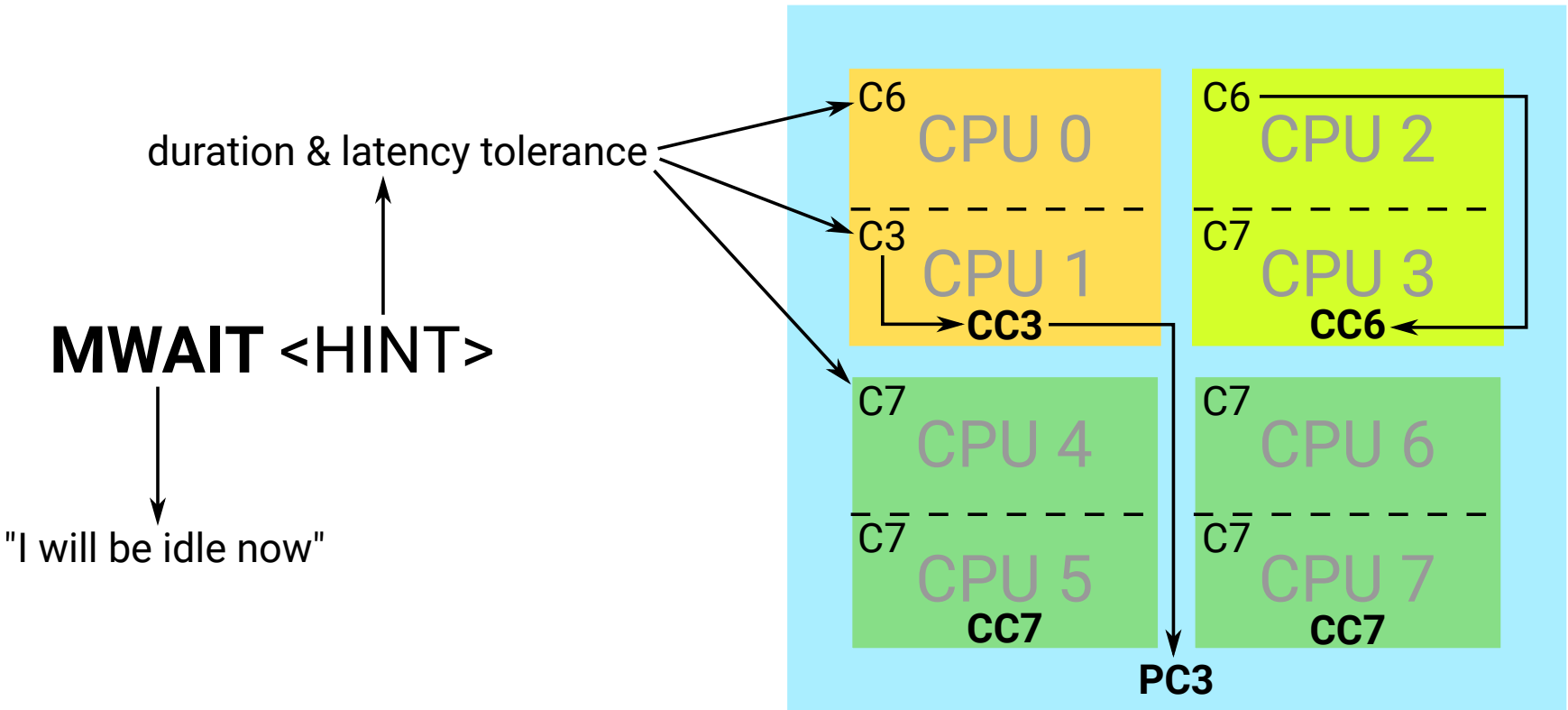
Interrupt Timing Measurements May Be Affected Too





Intel OpenSource TECHNOLOGY CENTER

Example: Core and Package C-states

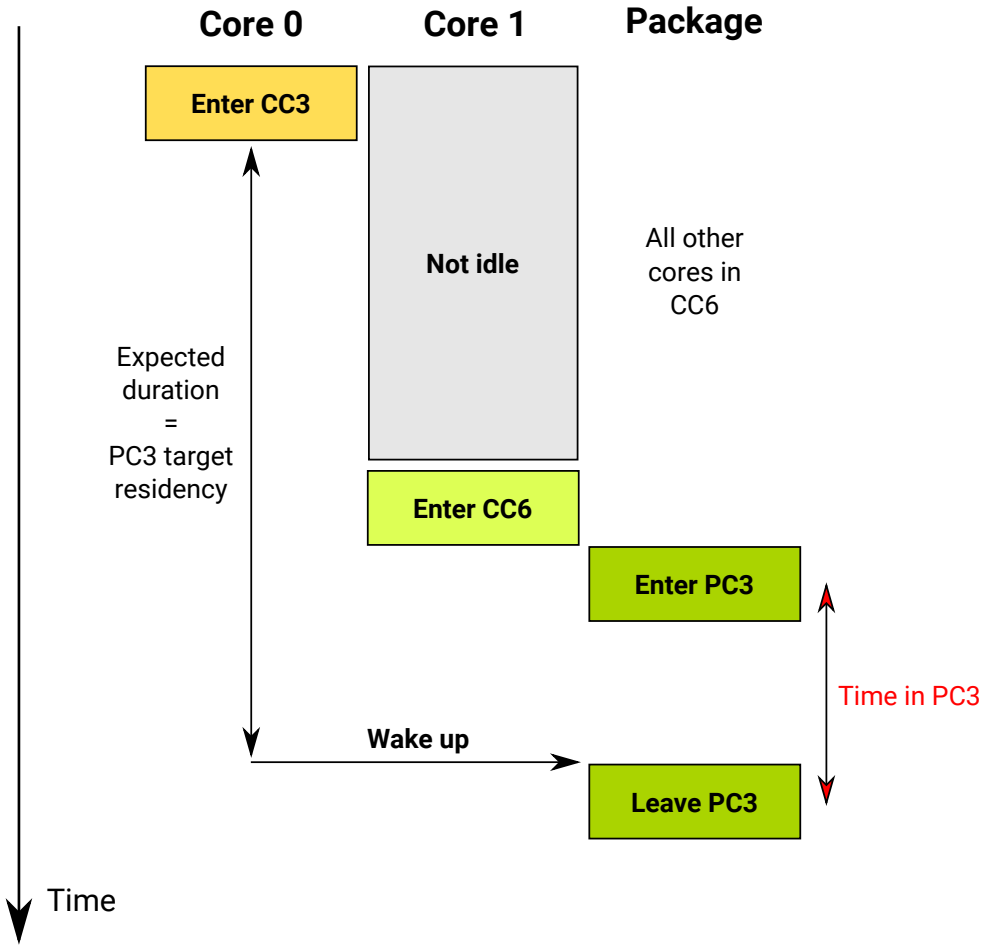


Intel OpenSource TECHNOLOGY CENTER

What Target Residency To Use for a "Combo" State?

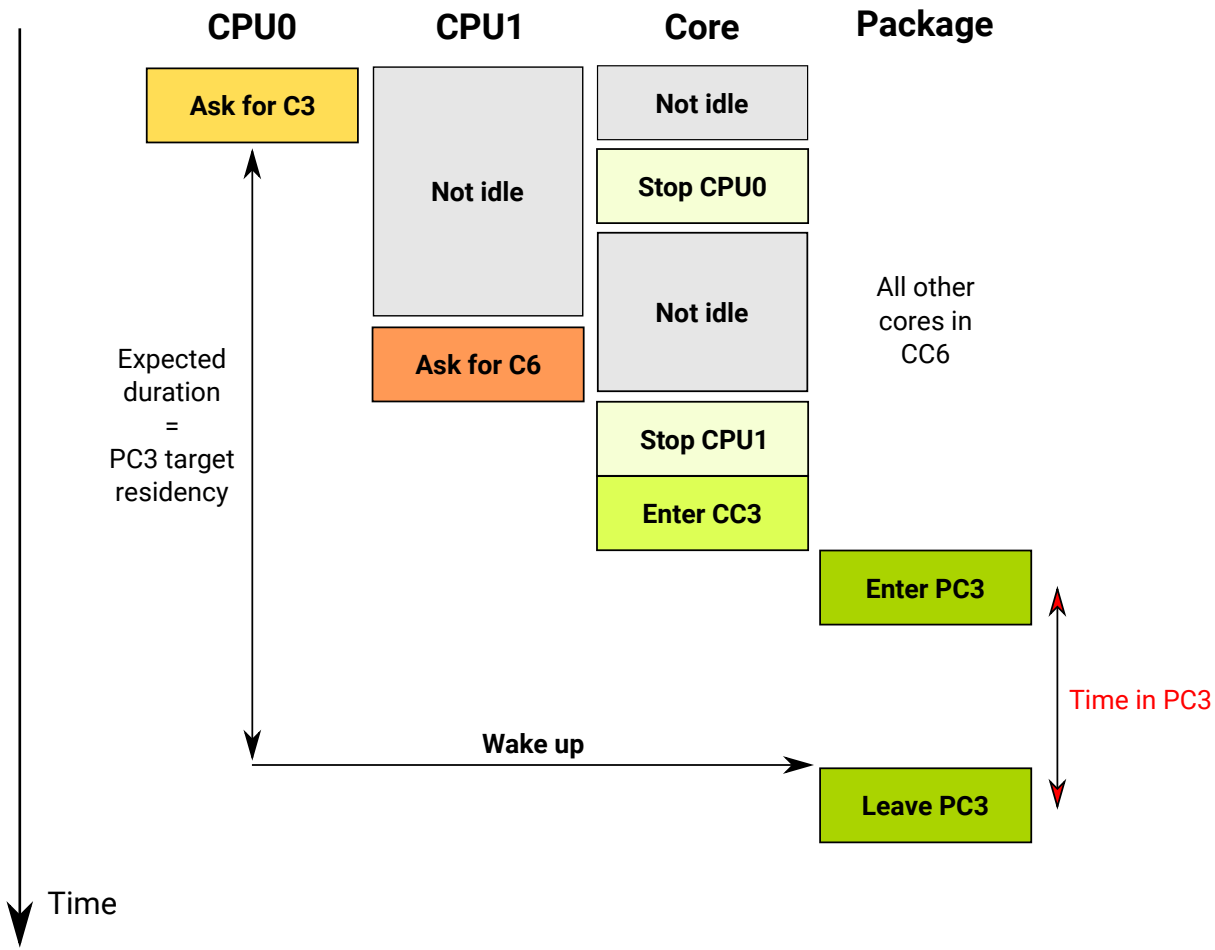
Intel OpenSource TECHNOLOGY CENTER

Target Residencies May Be Missed Due to Interactions (1)



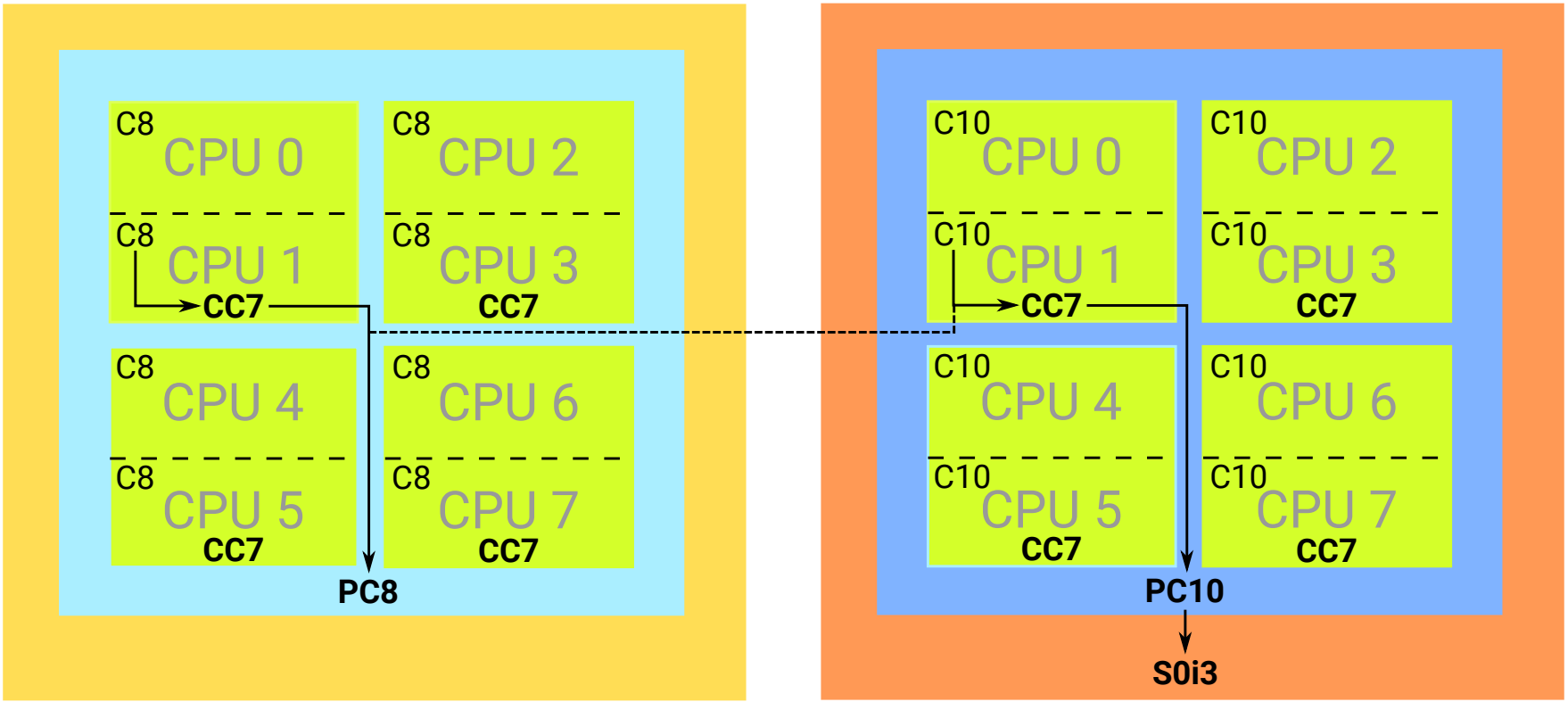
Intel OpenSource TECHNOLOGY CENTER

Target Residencies May Be Missed Due to Interactions (2)



Intel OpenSource TECHNOLOGY CENTER

Overlapping Hints and Missed Opportunities



Intel OpenSource TECHNOLOGY CENTER

Fall-through Idle States Idea

Assume target residency (TR) of the next idle state

Selected for latency reasons only.

Results (12 runs on Dell XPS13 9360, TEO governor)

Metric	**Baseline**	**C6 – C10: TR of C6**	**Δ**	**%**
BaseMark	301.0 ± 6.5	305.9 ± 5.6	4.9	1.7
JetStream2 (1)	54.5 ± 1.7	55.1 ± 1.6	0.6	1.1
speedometer	44.9 ± 0.4	45.2 ± 0.3	0.3	0.7
JetStream2 (2)	53.5 ± 1.4	54.4 ± 1.6	0.9	1.6
Active Watt (pkg)	13.1 ± 0.1	13.1 ± 0.2	0.0	0.0
Idle Watt (pkg)	0.93 ± 0.08	0.88 ± 0.07	−0.05	−5.8

Intel OpenSource TECHNOLOGY CENTER

Questions?

Intel OpenSource TECHNOLOGY CENTER

References


Marta Rybczyńska, *Improving idle behavior in tickless systems* (https://lwn.net/Articles/775618/).


Tom Yates, *What's a CPU to do when it has nothing to do?* (https://lwn.net/Articles/767630/).


Rafael J. Wysocki, *CPU Idle Loop Rework* (https://kernel-recipes.org/en/2018/talks/cpu-idle-loop-rework/).


T. Ilsche, M. Hähnel, R. Schöhne, M. Bielert, D. Hackenberg,
Powernightmares: The Challenge of Efficiently Using Sleep States on Multi-Core Systems
(https://tu-dresden.de/zih/forschung/ressourcen/dateien/projekte/haec/powernightmares.pdf?lang=en).


Rafael J. Wysocki, *Power Management Challenges in Linux* (https://www.linuxplumbersconf.org/2017/ocw//system/presentations/4652/original/linux_pm_challenges.pdf).


Rafael J. Wysocki, *PM Infrastructure in the Linux Kernel – Current Status and Future*
(https://events.linuxfoundation.org/sites/events/files/slides/kernel_PM_infra_0.pdf).


Jonathan Corbet, *The cpuidle subsystem* (http://lwn.net/Articles/384146/).

Intel OpenSource TECHNOLOGY CENTER

Disclaimer

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at `www.intel.com`.

Intel OpenSource TECHNOLOGY CENTER