
THE LINUX FOUNDATION
OPEN SOURCE SUMMIT
EUROPE
Balancing Power Management and Real-Time Performance on Intel Platforms
Junxiao Chang, Intel

Agenda

- Real-Time Performance Fundamentals
- Intel Power Management Overview
- CPU Frequency and Turbo Mode Impact
- CPU Idle Driver Impact
- CPU Consolidation Optimization
- Power vs Latency Trade-off Analysis
- Best Practice Recommendations

Real Time Performance

- Determinism: Ensures high-priority tasks are scheduled within predictable timing boundaries.
- Low Jitter: Minimizes execution-time variation to ensure deterministic responses.
- Maximum scheduling latency is the primary metric for evaluating real-time performance.

Real-Time Best Known Configuration (BKC)

- PREEMPT_RT alone is insufficient. Additional system tuning is required to achieve optimal real-time performance.
- Performance optimization configurations include:

RT Optimization	Purpose
CPU Isolation	Reduce interference
CPU Idle Tuning	Minimize wake-up latency
CPU Frequency Tuning	Stabilize execution timing
Timer Reduction	Lower scheduling overhead
Task Offloading	Reduce OS noise

Real Time Best Known Configuration Example

Kernel command line parameters

- clocksource=tsc
- idle=poll
- intel_pstate=disable
- isolcpus=2-3
- noefi
- irqaffinity=0
- ...

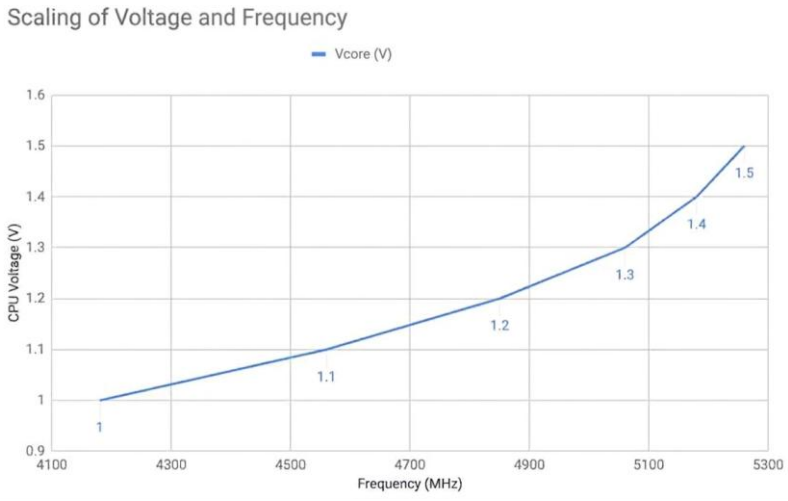
Kernel config options

- CONFIG_PREEMPT_RT=y
- CONFIG_CPU_ISOLATION=y
- CONFIG_NO_HZ_FULL=y
- CONFIG_HIGH_RES_TIMERS=y
- ...

Note: These BKC could be found in Intel TCC user guide or real time tuning document in linux distributions website.

CPU Power Consumption

- CPU power consumption increases directly with higher clock frequency and voltage.
- Power consumption follows the formula $P = C \times V^2 \times f$, meaning dynamic power scales linearly with frequency but has a squared relationship with voltage.
- Roughly linear relationship between V and f - Higher frequency requires higher voltage, causing power to increase disproportionately ($P \propto V^2 f$).



Power Management & Power Meter Tool “turbostat”

- Electric power or current drain (in case of constant voltage) can be used for power management benchmark.
- The Linux **turbostat** utility reports Intel processor power and performance metrics. Its source is in “tools/power/x86/turbostat/” in Linux kernel tree
- “PkgWatt” in turbostat output reports total processor package power consumption.

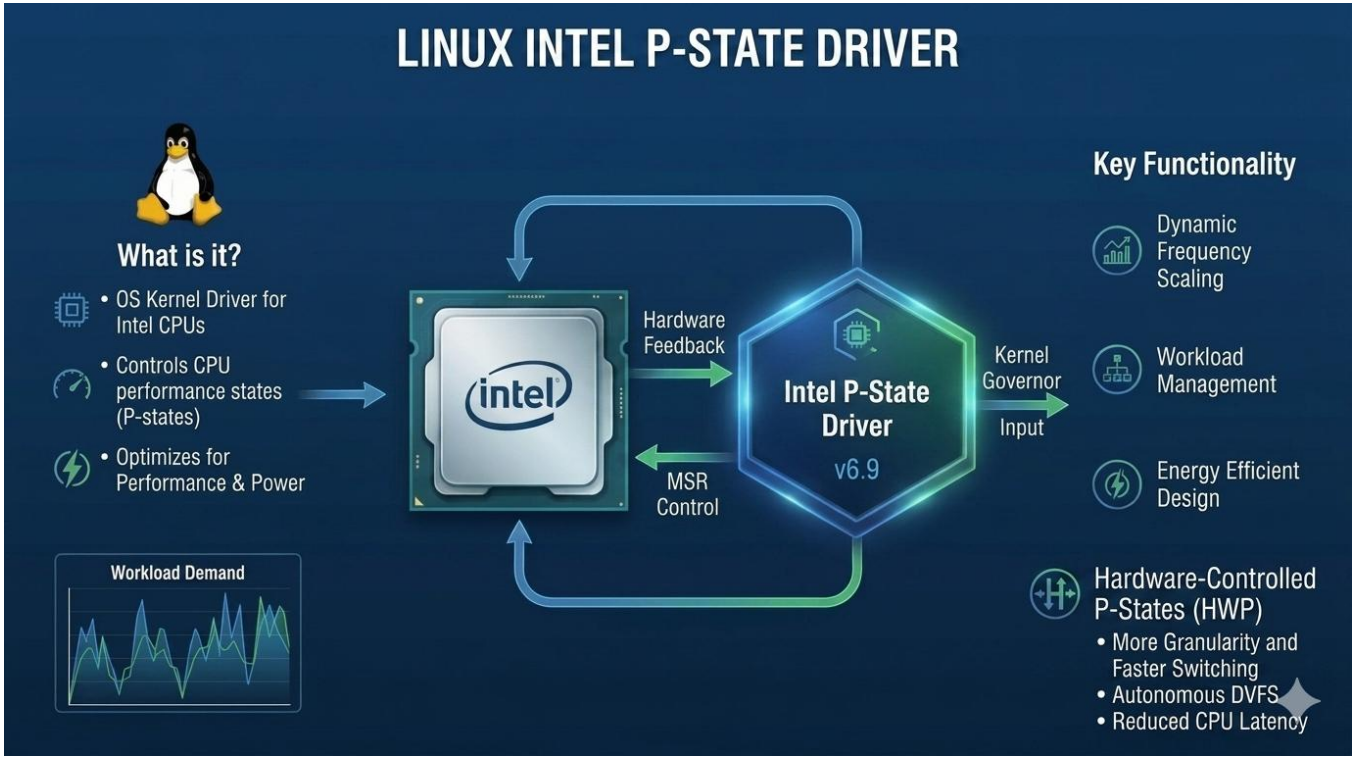
```
cpu0: MSR_IA32_PACKAGE_THERM_INTERRUPT: 0x02000003 (100 C, 100 C)
cpu0: MSR_MISC_FEATURE_CONTROL: 0x00000000 (L2-Prefetch L2-Prefetch-pair L1-Prefetch L1-IP-Prefetch)
Core    CPU     Avg_MHz Busy%   Bzy_MHz TSC_MHz IPC             IRQ             NMI     SMI     LLCMRPS LLC%hit L2MRPS  L2%hit  CPU%c1  CPU%c6  CPU%c7  CoreTmp CoreThr PkgTmp  Totl%C0 Any%C0  GFX%C0  CPUGFX% Pkg%pc2 Pkg%pc6 Pk%pc10 CPU%LPI SYS%LPI PkgWatt CorWatt GFX
-       -       3417    99.77   3424    3686    0.11    12413048        0       0       1       81.47   2       86.89   0.00    0.00    0.00    58      0       58      1594.12 99.63   0.01    0.01    0.00    0.00    0.00    0.00    0.00    0.00    23.63   22.67   0.0
0       0       4288    99.77   4298    3686    0.14    617312          0       0       0       70.39   0       63.34   0.00    0.00    0.00    55      0       58      1594.12 99.63   0.01    0.01    0.00    0.00    0.00    0.00    0.00    0.00    23.63   22.67   0.0
4       1       4289    99.77   4299    3686    0.14    605623          0       0       0       89.13   0       62.89   0.00    0.00    0.00    55      0
8       2       4280    99.77   4290    3686    0.14    3064954         0       0       0       95.83   0       91.99   0.00    0.00    0.00    58      0
12      3       4290    99.77   4300    3686    0.14    605022          0       0       0       99.50   0       5.08    0.00    0.00    0.00    54      0
16      4       3193    99.77   3200    3686    0.10    620148          0       0       0       72.36   0       86.11   0.00    0.00    0.00    56      0
17      5       3193    99.77   3200    3686    0.10    619703          0       0       0       74.72   0       88.47   0.00    0.00    0.00    56      0
18      6       3193    99.77   3200    3686    0.10    621844          0       0       0       75.23   0       89.89   0.00    0.00    0.00    56      0
19      7       3193    99.77   3200    3686    0.10    622440          0       0       0       64.60   0       81.32   0.00    0.00    0.00    56      0
20      8       3193    99.77   3200    3686    0.10    652892          0       0       0       62.87   0       90.19   0.00    0.00    0.00    55      0
21      9       3193    99.77   3200    3686    0.10    654652          0       0       0       75.35   0       92.08   0.00    0.00    0.00    55      0
22      10      3193    99.77   3200    3686    0.10    654564          0       0       0       64.71   0       91.75   0.00    0.00    0.00    55      0
23      11      3193    99.77   3200    3686    0.10    650127          0       0       0       78.15   0       92.58   0.00    0.00    0.00    55      0
```

Best RT Performance and Best Power Consumption

	Configuration	PM first	Latency first
CPU Online	All Online	All Online	All Online
CPU Idle	intel_idle		
CPU Freq	intel pstate		
Turbo Mode	Enabled		
	Latency:	19us	2us
	Power Wattage:	0.8W	23.6W

Hardware: PTL Ultra X7 358H, DDR5-6400 16M RAM. **Software**: Ubuntu 26.04 + v7.0.1-rt2 Kernel

Intel CPU Frequency Driver



Intel CPU Frequency Turbo Mode

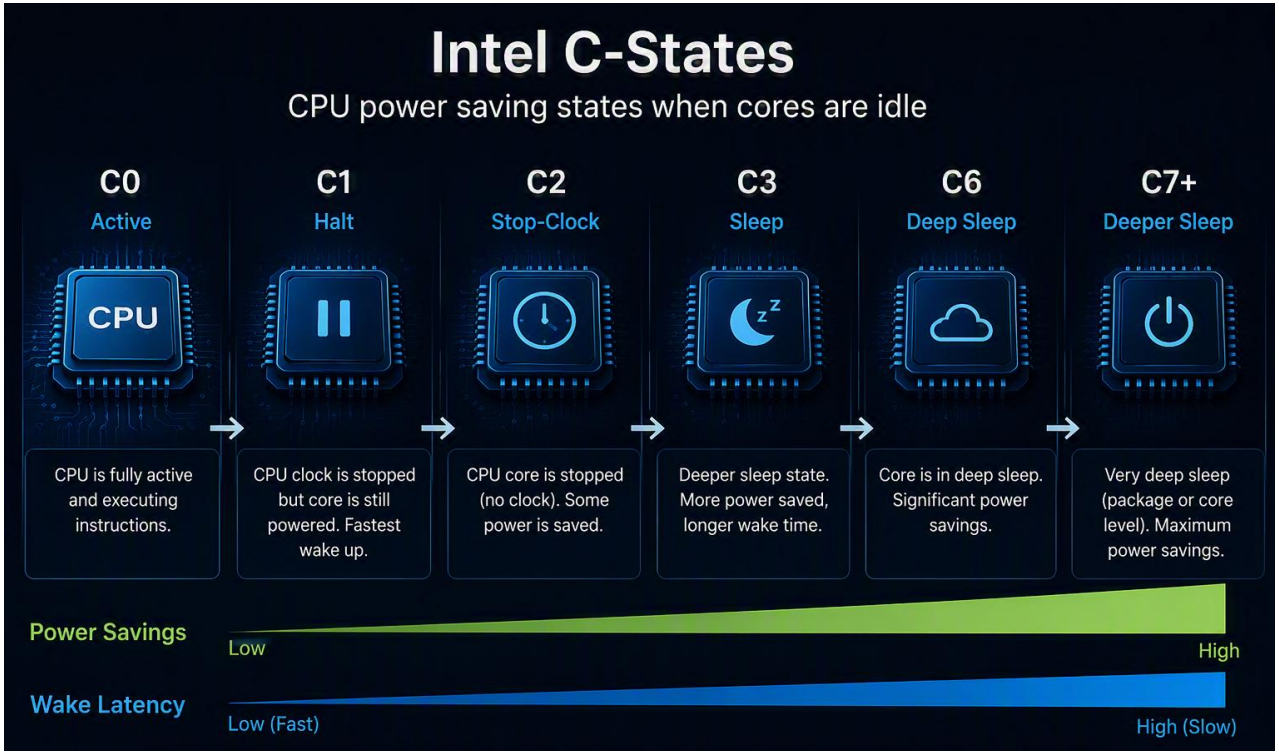
- Processor can operate at its base frequency under nominal load.
- When thermal and power headroom are available, frequency can increase beyond the base frequency.
- Turbo Boost improves peak performance while maintaining platform safety limits.
- Turbo operation is automatically managed by hardware and OS power policies.

Enable CPU P-state Frequency Driver

	Configuration	Turbo On	Turbo Off
CPU Online	All cores are online		
CPU Idle	intel_idle		
CPU Freq	intel pstate		
Turbo Mode	Enabled		
	Latency:	3us	4us
	Power Wattage:	24.2W	5.2W

Turbo improves latency by only 1µs while increasing power by nearly 5×.

Intel CPU Idle Driver



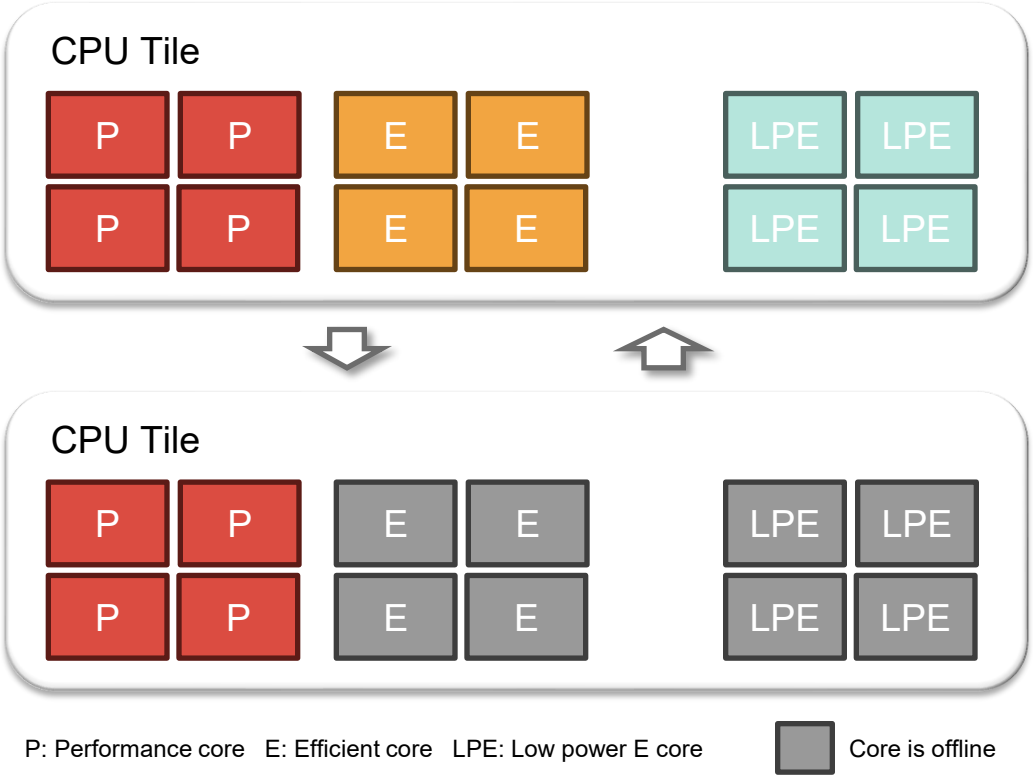
Enabling CPU Idle Driver

	Configuration	Idle On	Only C1 On	Idle Off
CPU Online	All cores are online			
CPU Idle	intel_idle	 All C states are enabled	 Only C1 is enabled	
CPU Freq	intel pstate			
Turbo Mode	Enabled			
	Latency:	23us	16us	2us
	Power Wattage:	0.8W	1.2W	23.6W

CPU Idle is the largest contributor to power savings and latency degradation.

CPU Consolidation

- Dynamically enable or disable CPU cores based on workload demands.
  Note: CPU can be disabled by:
  "echo 0 > /sys/devices/system/cpu/cpuX/online"
- Apply it only when all system functions are operating normally and under control.



Offline All E and LPE Cores

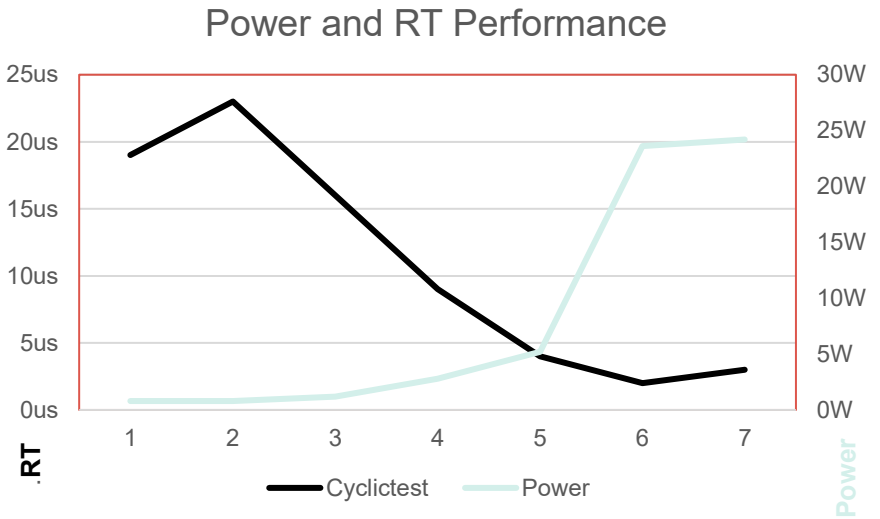
	Configuration	All Online	E & LPE Off
CPU Online		All Online	E/LPE Offline
CPU Idle	intel_idle		
CPU Freq	intel pstate		
Turbo Mode	Enabled		
	Latency:	4us	9us
	Power Wattage:	5.2W	2.8W

Core consolidation can reduce power with moderate latency impact.

Power Performance In All Cases

	CPUIdle	CPUFreq	Turbo	Online	Cyclictest	Power
1	On	On	On	All	19	0.8
2	On	Off	Off	All	23	0.8
3	C1 On	Off	Off	All	16	1.2
4	Off	On	Off	P cores	9	2.8
5	Off	On	Off	All	4	5.2
6	Off	Off	Off	All	2	23.6
7	Off	On	On	All	3	24.2

Note: cyclictest unit is microsecond; power unit is watt.



Balancing Power And RT Performance

- Power-constrained systems, low power consumption is preferred – Enable CPU Idle, and CPU Freq. driver.

	CPUIdle	CPUFreq	Turbo	Online	Cyclictest	Power
1	On	On	On	All	19	0.8

- Balanced systems and 10us scheduling latency is acceptable – Disable CPU idle, enable CPU Freq but disable its turbo mode. Online/Offline the CPUs depends on system workload.

	CPUIdle	CPUFreq	Turbo	Online	Cyclictest	Power
4	Off	On	Off	P cores	9	2.8
5	Off	On	Off	All	4	5.2

- Hard real-time systems and scheduling latency is the top priority, disable CPU Idle and CPU Frequency scaling.

	CPUIdle	CPUFreq	Turbo	Online	Cyclictest	Power
6	Off	Off	Off	All	2	23.6

What's The Best Configuration?

There is no universal best configuration, optimize for the performance metrics that matter most to use case.

The optimal configuration depends on:

- Latency requirement
- Power budget
- Thermal constraints
- Workload characteristics
- Platform architecture

Key Takeaways

For Intel embedded platforms:

- CPU Idle has the largest impact on power.
- CPU Idle also has the largest impact on latency.
- CPU Frequency scaling provides a good trade-off mechanism.
- Platform-specific validation is essential.

Reference

- Public Intel® Time Coordinated Compute (TCC) User Guide
- Energy Analysis Metrics Reference
- intel_pstate CPU Performance Scaling Driver

Q/A