Changes in Linux PM Core Since v3.14

Rafael J. Wysocki

Intel Open Source Technology Center

July 21, 2015
Suspend-to-Idle Support

Suspend-to-Idle Idea

- **Suspend**
  - Call Notifiers
  - Freeze Tasks
    - Device Suspend
      - prepare()
      - suspend()
      - suspend_late()
      - suspend_noirq()
    - Nonboot CPU Offline
    - System Core Offline
    - Platform Offline
  - Nonboot CPU Online
  - System Core Online
  - Platform Online

- **Resume**
  - Call Notifiers
  - Thaw Tasks
    - Device Resume
      - complete()
      - resume()
      - resume_early()
      - resume_noirq()
    - Wait For a Wakeup Interrupt

- **Interrupt handlers run**
- **No interrupt handlers**

Rafael J. Wysocki (SSG OTC)
PM Changes Since v3.14
July 21, 2015
Suspend-to-Idle Support

Suspend-to-Idle: Why Bother?

Suspend-to-idle vs full suspend
- Platform support not required (not used).
- May be made faster.

Suspend-to-idle vs runtime idle
- Fewer interrupts (less interrupt jitter).
- Forced (rather than voluntary) transitions to low-power states.
Suspend-to-Idle Support

Suspend-to-Idle: Quiescent Mode (Shipped In 4.0)

Device Suspend

Is this the last non-idle CPU?

YES

Suspend Timekeeping

NO

Stop Local Tick

Disabling Interrupts!

Device Resume

Wake?

YES

Interrupt

NO

Device Suspend

Go To Idle

Is this the 1st non-idle CPU?

YES

Resume Timekeeping

NO

Start Local Tick

PM Changes Since v3.14
July 21, 2015
Suspend-to-Idle Support

Suspend-to-Idle: Wakeup

Device Suspend
- prepare()
- suspend()
- suspend_late()
- suspend_noirq()

Device Resume
- complete()
- resume()
- resume_early()
- resume_noirq()

Go To Idle

Interrupt

Wake?

IRQD_WAKEUP_ARMED == 1?

enable_irq_wake()

IRQD_WAKEUP_SET = 1

IRQD_WAKEUP_SET == 1

IRQD_WAKEUP_ARMED = 1

Runtime IRQ mode
Wakeup IRQ mode
Assorted Changes Related To System Suspend/Resume

- Asynchronous execution of all device suspend/resume callbacks (3.15).
- Delayed resuming of runtime-suspended devices during system suspend (3.15).
- New helpers to run device runtime PM callbacks during suspend/resume (3.15).
- PM core support for avoiding to resume runtime-suspended devices during system suspends (3.16).
- ACPI PM domain support for avoiding to resume runtime-suspended devices during system suspends (3.16).
- Suspend/resume trace events (3.16).
- Version 3.0 of analyze_suspend.py (3.17).
- Consolidation of device PM Kconfig options (CONFIG_PM_RUNTIME removal; 3.19).
- PCI core support for avoiding to resume runtime-suspended devices during system suspend (4.0).
- PM tracing support for the suspend phase of system suspend/resume transitions (4.1).
- Support for automated device wakeup IRQ handling (4.2-rc).
CPUfreq

Core changes

- Ton of fixes (each kernel version).
- Preserve sysfs directories/files over CPU offline/online (v4.2-rc).

intel_pstate

- New CPU IDs (each kernel version).
- Hardware Managed Performance States (HWP) support (3.19).
- “force” command line switch (3.19).
CPUfreq Contd.

**cpufreq-dt**
- Common driver for DT-based platforms.
- DT bindings for Operating Performance Points.
- Multiple platforms use it.