Advances in CPU Idle Time Management

Rafael J. Wysocki

Intel Open Source Technology Center

August 30, 2018
CPUs: Busy Vs Idle
CPU Idle States Properties

- **Power [a. u.]**
- **Time [a. u.]**
- **Entry**
- **Exit**
- **Residency**

Rafael J. Wysocki

CPU Idle @ OSS 2018
August 30, 2018
CPU Idle States Example: C-states on Intel Hardware

MWAIT <HINT>
High-level CPU Idle Time Management Control Flow

- CPU scheduler
  - NO
    - CPU is idle
  - YES
    - Estimate idle time
      - menu
      - ladder
    - Select state
      - Governor
      - intel_idle
      - ACPI idle
    - Enter state
Complication: CPU Scheduler Tick Timer
Original Idle Loop Design Issue

1. CPU scheduler
2. Try to stop tick
3. need_resched()?
   - YES: Restart tick
   - NO:
     - Estimate idle time
       - May be too short!
     - Select state
     - Governor
     - Enter state
     - Interrupt
Short Idle Duration Prediction Problem

<table>
<thead>
<tr>
<th>Actual</th>
<th>Long Idle</th>
<th>Short Idle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long Idle</td>
<td><strong>WIN</strong></td>
<td><strong>LOSS</strong></td>
</tr>
<tr>
<td>Short Idle</td>
<td><strong>LOSS</strong></td>
<td><strong>LOSS</strong></td>
</tr>
</tbody>
</table>
Issues With The menu Idle Governor
Redesigned Idle Loop (Linux* 4.17 And Later)

- CPU scheduler
- need_resched()?
  - NO
  - Select state
  - Estimate idle time
  - Governor
  - Restart tick
  - Tick stopped?
    - NO
    - Restart tick
    - Tick stopped?
      - YES
      - Try to stop tick
      - Stop tick or tick stopped?
        - NO
        - Enter state
        - Interrupt
        - YES
        - Try to stop tick
  - YES
  - Enter state
  - Tick stopped?
    - YES
Short Idle Duration Prediction Problem Mitigated

```
<table>
<thead>
<tr>
<th>Predicted</th>
<th>Long Idle</th>
<th>Short Idle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long Idle</td>
<td>WIN</td>
<td>LOSS</td>
</tr>
<tr>
<td>Short Idle</td>
<td>Neutral</td>
<td>WIN</td>
</tr>
</tbody>
</table>
```

Rafael J. Wysocki
Idle Power (Intel OTC Server Power Lab)
Idle Power (Intel OTC Server Power Lab) Continued
Idle Power (Intel OTC Server Power Lab) Continued
CPU Idle Time Management And Power Domains
Questions?
References


5. J. Corbet, *The cpuidle subsystem* (http://lwn.net/Articles/384146/).
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 and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries. *Other names and brands may be claimed as the property of others.

© Intel Corporation