Power Management

Presented by
Jaroslav Škarvada
jskarvad@redhat.com
Outline

- Introduction
- PM QoS
- PowerTOP 2.0
- pm-utils
- Manual tunings
- Future plans
Introduction

- Power Management (PM) SIG:
  - http://fedoraproject.org/wiki/SIGs/PowerManagement

- Goals:
  - Improving PM in Fedora.
  - Identification of problematic components.
  - Tuning for energy savings:
    - Longer operation from battery for mobile users.
    - Cost savings for desktops / servers.
Weekly meetings:
- Each Wednesday at 14:00 UTC on #fedora-meeting

Other communication channels:
- Mailing list: power-management
- IRC channel: #power

To do list:
- [http://fedoraproject.org/wiki/SIGs/PowerManagement/Todos](http://fedoraproject.org/wiki/SIGs/PowerManagement/Todos)

F15 test day planned on 2011-03-24
Bugzilla PM Trackers

- Bugzilla main PM tracker:
  - BZ#484668

- Sub-trackers:
  - For apps waking too much: BZ#204948
  - For disk-over-eager apps: BZ#454582
Power/Energy

- Slower doesn't mean energy / costs savings:
  - HP ProLiant DL360 G6, Xeon E5504 @ 2.00 GHz
  - # mock kernel-2.6.35.10-81.fc14.src.rpm

<table>
<thead>
<tr>
<th>CPU speed [GHz]</th>
<th>t [s]</th>
<th>P [W]</th>
<th>E [Wh]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0</td>
<td>4076</td>
<td>90.0</td>
<td>101.9</td>
</tr>
<tr>
<td>1.6</td>
<td>4699</td>
<td>82.1</td>
<td>107.2</td>
</tr>
</tbody>
</table>

- Lower $P \Rightarrow$ Better cooling, power capping.
- Lower $E \Rightarrow$ Longer operation from battery, costs savings.
Deeper sleep modes ⇒ increased latency

The kernel PM-QoS infrastructure tries to address this problem.

- User space applications signals its latency / throughput requirements to kernel.
- The kernel components uses best effort to serve the requirements.
  - Currently mac80211 stack and cpuidle support this.
Following objects are supported:
- cpu_dma_latency (usec)
- network_latency (usec)
- network_throughput (kB/s)

User space app register itself by opening associated file in /dev and writes there its request:
- The request is valid till the file handle is hold.
- $ exec 3> /dev/cpu_dma_latency;
  echo 999 >&3; ...; exit
PowerTOP 2.0 – I

- Author: Arjan van de Ven <arjan@linux.intel.com>
- Complete rewrite of 1.x codebase.

Features:
- Redesigned tab based user interface.
- Uses 'perf' infrastructure for improved accuracy.
- Tracks the power behaviour of various devices.
- Power estimation engine.
- Ability to create HTML reports.
Run from battery for power consumption tracking.

For best results calibrate the power estimation engine:

```bash
# powertop --calibrate
```

It will cycle through various settings (brightness, rfkill, CPU loads).

Generate HTML report:

```bash
# powertop --html
```
Overview:

The battery reports a discharge rate of 14.3 W
The estimated remaining time is 152 minutes

Summary: 230.4 wuks/second, 0.0 GPU ops/second and 0.0 VFS ops/sec

<table>
<thead>
<tr>
<th>Power est.</th>
<th>Usage</th>
<th>Events/s</th>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.09 W</td>
<td>2.3 ms/s</td>
<td>86.3</td>
<td>Interrupt</td>
<td>PS/2 Touchpad / Keyboard / Mo</td>
</tr>
<tr>
<td>692 mW</td>
<td>2571 rpm</td>
<td></td>
<td>Device</td>
<td>Laptop fan</td>
</tr>
<tr>
<td>629 mW</td>
<td>7.2 ms/s</td>
<td>48.9</td>
<td>Process</td>
<td>/usr/bin/Xorg :0 -nr -verbose gnome-terminal</td>
</tr>
<tr>
<td>442 mW</td>
<td>4.6 ms/s</td>
<td>34.4</td>
<td>Process</td>
<td>nautilus</td>
</tr>
<tr>
<td>415 mW</td>
<td>329.5 pkts/s</td>
<td></td>
<td>Device</td>
<td>Network interface: eth0 (e100</td>
</tr>
<tr>
<td>174 mW</td>
<td>3.2 ms/s</td>
<td>13.4</td>
<td>Process</td>
<td>nautilus</td>
</tr>
<tr>
<td>142 mW</td>
<td>92.0 µs/s</td>
<td>11.2</td>
<td>Interrupt</td>
<td>[41] i915</td>
</tr>
<tr>
<td>125 mW</td>
<td>100.0%</td>
<td></td>
<td>Device</td>
<td>Audio codec hwC0D0: Conexant</td>
</tr>
<tr>
<td>125 mW</td>
<td>100.0%</td>
<td></td>
<td>Device</td>
<td>Audio codec hwC0D3: Intel</td>
</tr>
<tr>
<td>64.1 mW</td>
<td>204.1 µs/s</td>
<td>5.1</td>
<td>Timer</td>
<td>tick_sched_timer</td>
</tr>
<tr>
<td>63.2 mW</td>
<td>85.7 µs/s</td>
<td>5.0</td>
<td>Interrupt</td>
<td>[6] tasklet(softirq)</td>
</tr>
<tr>
<td>52.5 mW</td>
<td>58.4 µs/s</td>
<td>4.2</td>
<td>Timer</td>
<td>hrtimer_wakeup</td>
</tr>
<tr>
<td>34.8 mW</td>
<td>75.5 µs/s</td>
<td>2.8</td>
<td>Process</td>
<td>[scsi eh 1]</td>
</tr>
<tr>
<td>28.5 mW</td>
<td>80.9 µs/s</td>
<td>2.3</td>
<td>Process</td>
<td>udisks-daemon: polling /dev/s</td>
</tr>
<tr>
<td>15.2 mW</td>
<td>51.6 µs/s</td>
<td>1.2</td>
<td>Interrupt</td>
<td>[7] sched(softirq)</td>
</tr>
</tbody>
</table>

ESC Exit
## Tunables:

<table>
<thead>
<tr>
<th>PowerTOP 1.97</th>
<th>Overview</th>
<th>Idle stats</th>
<th>Frequency stats</th>
<th>Device stats</th>
<th>Tunables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bad</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bad</td>
<td>Wireless Power Saving for interface wlan0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bad</td>
<td>Enable SATA link power management for /dev/sda</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bad</td>
<td>VM writeback timeout</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bad</td>
<td>Enable Audio codec power management</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bad</td>
<td>Autosuspend for USB device ThinkPad Bluetooth with Enhanced Data Rate II (Le</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>NMI watchdog should be turned off</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Power Aware CPU scheduler</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Bluetooth device interface status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Autosuspend for USB device UHCI Host Controller [usb8]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Autosuspend for USB device UHCI Host Controller [usb6]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Autosuspend for USB device UHCI Host Controller [usb7]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Autosuspend for USB device EHCI Host Controller [usb1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Autosuspend for USB device EHCI Host Controller [usb2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Autosuspend for USB device UHCI Host Controller [usb5]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Autosuspend for USB device UHCI Host Controller [usb3]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Autosuspend for USB device Chicony Electronics Co., Ltd. [1-6]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Autosuspend for USB device UHCI Host Controller [usb4]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Wake-on-lan status for device eth0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Wake-on-lan status for device vboxnet0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<ESC> Exit | <Enter> Toggle tunable
Improved bug reporting helper script.

If you encounter problem with suspend / hibernate:

- Try to run pm-suspend / pm-hibernate from console.
- If you have problem with you video card, try with various video quirks:
  - # pm-suspend --quirk-test YOUR QUIRKS
- List of quirks can be obtained by:
  - # pm-suspend --help
  - http://hal.freedesktop.org/quirk/quirk-suspend-explain.html
pm-utils – II

- Also try to by-pass the pm-utils:
  - `# echo mem > /sys/power/state`
  - `# echo disk > /sys/power/state`

- File a bug.

- Run bug report info script:
  - `# pm-utils-bugreport-info.sh`
  - Attach its output to the bugzilla.
Run scripts on AC / battery change:

- Put them into: `/etc/power.d`
- Each will be run with $1 parameter:
  - true – system was switched to battery,  
  - false – system was switched to AC.

Override default config, e.g. disable the `01grub` hook:

```
# echo 'HOOK_BLACKLIST="01grub"' > /etc/pm/config.d/user
```
Since tuned-0.2.19 it supports EEEPC FSB downclocking:

- Reduce FSB:
  - `# tuned-adm profile laptop-battery-powersave`

- Restore FSB:
  - `# tuned-adm profile default`

Asus 1000H:

<table>
<thead>
<tr>
<th>FSB</th>
<th>Pavgidle [W]</th>
<th>Pload [W]</th>
<th># of transactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>9.5</td>
<td>18.7</td>
<td>5597</td>
</tr>
<tr>
<td>Reduced</td>
<td>8.3</td>
<td>14.8</td>
<td>4330</td>
</tr>
</tbody>
</table>
Manual tunings – I

- For cases when “safe” defaults are not enough.
- Actual possibilities depends on HW used.
- ALPM max powersave (~ 1 W):
  - `echo min_power > /sys/class/scsi_host/host0/link_power_management_policy`
- HDD max powersave, spindown in 10 minutes (~ 0.9 W):
  - `hdparm -B 1 -S 120 /dev/sda`
- Disable bluetooth:
  - `# rfkill block bluetooth`
Switch ethernet to 100 Mbit (~ 0.3 W):
  - `ethtool -s eth0 advertise 0x0f`

Force max PS-poll for some Intel WiFi cards:
  - `iwpriv eth1 set_power 5`
  - `for i in /sys/bus/pci/devices/*/power_level ; do echo 5 > $i ; done`

Enable USB autosuspend (~ 0.3 W depends on dev):
  - `# for i in /sys/bus/usb/devices/*/power; do echo 1 > $i/autosuspend && echo auto > $i/control; done`
Future Plans

- Add PM support to services, e.g. support for battery / AC profiles.
- Add PM support to cronie (wrapper) to be able to postpone running of battery demanding tasks.
- Add support for XHCI (USB3) suspend.
- Better integration of PM-QoS (e.g. for PS-Poll).
- Add support for DDC/CI to control settings of external LCDs.
- Tuned profiles from the PowerTOP suggestions.

Power Management, Jaroslav Škarvada
References

- PM SIG:
  - http://fedoraproject.org/wiki/SIGs/PowerManagement

- Intel's Lesswatts.org:
  - http://www.lesswatts.org/

- tuned:
  - https://fedorahosted.org/tuned/

- pm-utils:
  - http://pm-utils.freedesktop.org
Questions?

Thank you.