Power Management

Presented by

Jaroslav Škarvada

jskarvad@redhat.com

redhat.

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.

PM SIG

- 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
- 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`

CPU speed [GHz]	t [s]	P [W]	E [Wh]
2.0	4076	90.0	**101.9**
1.6	4699	82.1	**107.2**

  - Lower $P \Rightarrow$ Better cooling, power capping.

  - Lower $E \Rightarrow$ Longer operation from battery, costs savings.

PM-QoS

- 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.

PM-QoS – usage

- 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.

PowerTOP 2.0 – II

- Run from battery for power consumption tracking.

- For best results calibrate the power estimation engine:

  - `# powertop --calibrate`

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

- Generate HTML report:

  - `# powertop --html`

PowerTOP 2.0 – III

- Overview:

```
PowerTOP 1.97      Overview   Idle stats   Frequency stats   Device stats   Tunables

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

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

Power est.        Usage        Events/s    Category    Description
   1.09 W        2.3 ms/s       86.3        Interrupt   PS/2 Touchpad / Keyboard / Mo
   692 mW       2571 rpm                    Device      Laptop fan
   629 mW        7.2 ms/s       48.9        Process     /usr/bin/Xorg :0 -nr -verbose
   442 mW        4.6 ms/s       34.4        Process     gnome-terminal
   415 mW      329.5 pkts/s                 Device      Network interface: eth0 (e100
   174 mW        3.2 ms/s       13.4        Process     nautilus
   142 mW       92.0 µs/s       11.2        Interrupt   [41] i915
   125 mW      100.0%                       Device      Audio codec hwC0D0: Conexant
   125 mW      100.0%                       Device      Audio codec hwC0D3: Intel
  64.1 mW      204.1 µs/s        5.1        Timer       tick_sched_timer
  63.2 mW       85.7 µs/s        5.0        Interrupt   [6] tasklet(softirq)
  52.5 mW       58.4 µs/s        4.2        Timer       hrtimer_wakeup
  34.8 mW       75.5 µs/s        2.8        Process     [scsi_eh_1]
  28.5 mW       80.9 µs/s        2.3        Process     udisks-daemon: polling /dev/s
  15.2 mW       51.6 µs/s        1.2        Interrupt   [7] sched(softirq)

<ESC> Exit |
```

PowerTOP 2.0 – IV

- Tunables:

```
PowerTOP 1.97      Overview   Idle stats   Frequency stats   Device stats   Tunables

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

<ESC> Exit | <Enter> Toggle tunable
```

pm-utils – I

- 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.

pm-utils – III

- 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`

tuned

- 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:

FSB	*Pavgidle* [W]	*Pload* [W]	# of transactions
Normal	9.5	18.7	**5597**
Reduced	8.3	14.8	**4330**

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
    ```

Manual tunings – II

- 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.

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.