Linux Suspend/Resume...
...at the Speed of Light

Len Brown, Principal Engineer,
Intel Open Source Technology Center

5-Oct, 2015
LinuxCon Europe
Dublin, Ireland
Acknowledgements

Todd Brandt – analyze_suspend.py maintainer
Rafael Wysocki – suspend/resume maintainer
Agenda

Concepts

Tools

Results

Future
Linux Suspend Types

$ cat /sys/power/state
disk mem standby freeze

Power savings

Speed
Saving Power with System Suspend

Power

Busy

Active Idle

System Suspend to Low-Power Idle

freeze

System Suspend to ACPI S3

cmem

5-Oct, 2015 LinuxCon Europe, Dublin, Ireland
## Suspend Trade-Offs

<table>
<thead>
<tr>
<th></th>
<th>Busy</th>
<th>Active Idle</th>
<th>Suspend to Low-Power Idle</th>
<th>Suspend to ACPI S3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Power</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Latency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Driver API</strong></td>
<td>Run-Time</td>
<td>Run-Time</td>
<td>System</td>
<td>System</td>
</tr>
<tr>
<td></td>
<td>Suspend</td>
<td>Suspend</td>
<td>Suspen</td>
<td>Suspen</td>
</tr>
<tr>
<td><strong>Display</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Network</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Interactive Laptop Scenario

- Wake
- Busy
- Display
- Resume
- Active Idle
- Sleep
- Suspend
Interactive Handheld Scenario

Power

Time

Wake
Resume
Busy
Display
Idle
Suspend
Display

5-Oct, 2015
LinuxCon Europe, Dublin, Ireland
Dark Resume Scenario

Power

Packet

Resume  Busy  Suspend

Time

5-Oct, 2015
LinuxCon Europe, Dublin, Ireland
Dark Resume Power vs Time

Assume:
- Active Power = 10x Suspend Power
- Suspend Time = 10x Active Time

So:
- Total Energy per interval = 20E
- Active Energy = 10/20 = 50% total energy
Assume:

- Active Power = 10x Suspend Power
- Suspend Time = 100x Active Time

So:

Total Energy per interval = 11E

Active Energy = 1/11 = 9% total energy
Dark Resume Challenge

Suspend + Resume time may exceed Active time
  Suspended battery life depends directly on suspend & resume performance
Resume Latency = packet latency
Going Faster

Start → Work → Finish

Time
Going Faster (strategy 1)

Less work, or less waiting

Before:

Start \[\rightarrow\] Finish

After:

Start \[\rightarrow\] Finish
Going Faster (strategy 2)

Same work, in parallel, but still synchronous

Before:

Start → Finish

After:

Start → Finish
Going Faster (strategy 3)

Same work, asynchronous

Before:

Start → Finish

After:

Start → Finish

Time

5-Oct, 2015 LinuxCon Europe, Dublin, Ireland
Going Faster (strategy 4)

Avoid work entirely

Before:

Start → Finish

After:

Start → Finish

Time
Agenda

Concepts

Tools

Results

Future
Measuring Suspend Speed

Method 1: Use external measuring device
Measuring Suspend Speed

Method 2: Boot with “initcall_debug”, examine msgbuf

```bash
$ dmesg | grep call
...
[ 661.392498] calling phy0+ @ 2367, parent: 0000:07:00.0
[ 661.417798] call phy0+ returned 0 after 24721 usecs
```
Measuring Suspend Speed

Method 3: Run `analyze_suspend`
Download and run `analyze_suspend.py`

```bash
$ git clone https://github.com/01org/suspendresume.git
$ cd suspendresume
$ sudo ./analyze_suspend.py
```

Generates output files in subdirectory: `suspend-yyyymmdd-HHMMSS`

- HTML output: `<hostname>_<mode>.html`
- raw dmesg output: `<hostname>_<mode>_dmesg.txt`
- raw ftrace output: `<hostname>_<mode>_ftrace.txt`

```bash
$ firefox suspend*/*.html
```
**analyze_suspend.py -h**

New script can re-analyze output of previous measurement

“initcall_debug” and dmesg used up through Linux 3.15, ftrace there-after

```
[general]
...
-m mode Mode to initiate for suspend ['freeze', 'mem', 'disk'] (default: mem)
-rtcwake t Use rtcwake to autoresume after <t> seconds (default: disabled)
...
-addlogs Add the dmesg and ftrace logs to the html output

[advanced]
...
-f Use ftrace to create device callgraphs (default: disabled)
...

[utilities]
...
[re-analyze data from previous runs]
-ftrace ftracefile Create HTML output using ftrace input
-dmesg dmesgfile Create HTML output using dmesg (not needed for kernel >= 3.15)
-summary directory Create a summary of all test in this dir
```
Agenda

Concepts

Tools

Results

Future
Suspend/Resume (mem)

ACPI S3: Firmware resume = 340ms
Display on: i915 resume > 1200ms
How to force Dark Suspend/Resume

Run-time suspend display before system-suspend:

```
$ xset -display :0 dpms force off
$ sleep 2
$ sudo analyze_suspend.py
```

Display will not be resumed upon system-resume, but availability is platform dependent...
Dark Suspend/Resume (mem)

ACPI S3: Firmware resume = 340ms
Display OFF
Suspend/Resume (freeze)

Firmware resume = 0
Display on: i915 resume > 1200ms
Dark Suspend/Resume (freeze)

Firmware resume = 0
Display OFF
analyze_suspend -f
Captures full ftrace call graph, parses in HTML GUI
{HTML file size ~ 64MB}
Things can go very wrong

Linux-4.0 sensors regression – no workaround
Things can go very wrong

pcieport resume 2900ms (https://bugzilla.kernel.org/show_bug.cgi?id=99751)
Workaround: boot with “pcie_ports=compat”
Things can go very wrong

850ms Serio1/psmouse due to lock contention in resume_complete. Fixed in Linux 4.2-rc1. Workaround: boot with “no_console_suspend”
Regarding the “speed of light”

O(25ms) to suspend and wake on this stripped-down Core2 desktop
UP, no sync, no GFX, serial console, no network, no USB, SSD drive, yes ACPI, but not FPDT

<table>
<thead>
<tr>
<th>Device Detail</th>
<th>ZOOM IN</th>
<th>ZOOM OUT</th>
<th>ZOOM 1:1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

sd @ 0:0:0:0 (Total Suspend: 5.454 ms Total Resume: 0.002 ms)

5.454 ms
suspend
Agenda

- Concepts
- Tools
- Results
- Future
What needs to be done?

Run analyze_suspend on more systems – help us!
  Prevent regressions
  Discover, report, fix more issues

Display, USB, Network, Audio
  When run-time suspended, stay suspended
  When must resume, go asynchronous
  Optimize actual resume latency

Wireless network re-association speed
Q & A
Linux “freeze” History

Linux-3.18:
  functional for 1st time, including wakeup

Linux-4.0:
  freeze timers, improves deep idle-state residency
Key Patches

ATA drives can take multiple SECONDS to resume
This patch makes that ASYNCHRONOUS, not blocking the resume path to user-space

In Linux v3.15-rc1:

commit 200421a80f6e0a9e39d698944cc35cba103eb6ce
Author: Todd Brandt <todd.e.brandt@linux.intel.com>
Date:   Fri Mar 14 13:52:54 2014 -0700

libata: async resume
Key Patches

Fix race condition in resume_complete
(boot with “no_console_suspend” may workaround)

In Linux v4.2-rc1:

commit 32e8d689dc12e29fcb6ba9c65a33473d0cbdfec8
Author: Todd E Brandt <todd.e.brandt@linux.intel.com>
Date:   Thu May 28 12:55:53 2015 -0700

PM / sleep: trace_device_pm_callback coverage in dpm_prepare/complete
Key Patches

Sensors regression – run-time vs system-suspend conflict
No workaround

**Linux 4.0 regression**

**Fixed in Linux v4.2, v4.1.4; NOT fixed in Linux 4.0-stable**

```
4.2-rc3 commit 1e25aa9641e8f3fa39cd5e46b4afcafd7f12a44b
4.2-rc4 commit 88cc7b4eeel7b9bcala64daee5adaa044cf72312
4.1.4 commit be43d21df90d10f5f10252c114f5fb024b7ba5ae
```

Author: Srinivas Pandruvada <srinivas.pandruvada@linux.intel.com>
Date: Mon Jun 1 16:36:27 2015 -0700

hid-sensor: Fix suspend/resume delay

[https://bugzilla.kernel.org/show_bug.cgi?id=102891](https://bugzilla.kernel.org/show_bug.cgi?id=102891)