

Linux Suspend/Resume

Len Brown, Rafael Wysocki, Todd Brandt - Intel



Rafael Wysocki

Linux Kernel PM Maintainer



Todd Brandt

Sleepgraph Maintainer



Today's Session

1. Level Set
2. sleepgraph tutorial
3. Discussion / Q&A Topics



Part 1 - Level Set



Linux System Suspend Types

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

Power savings



Speed

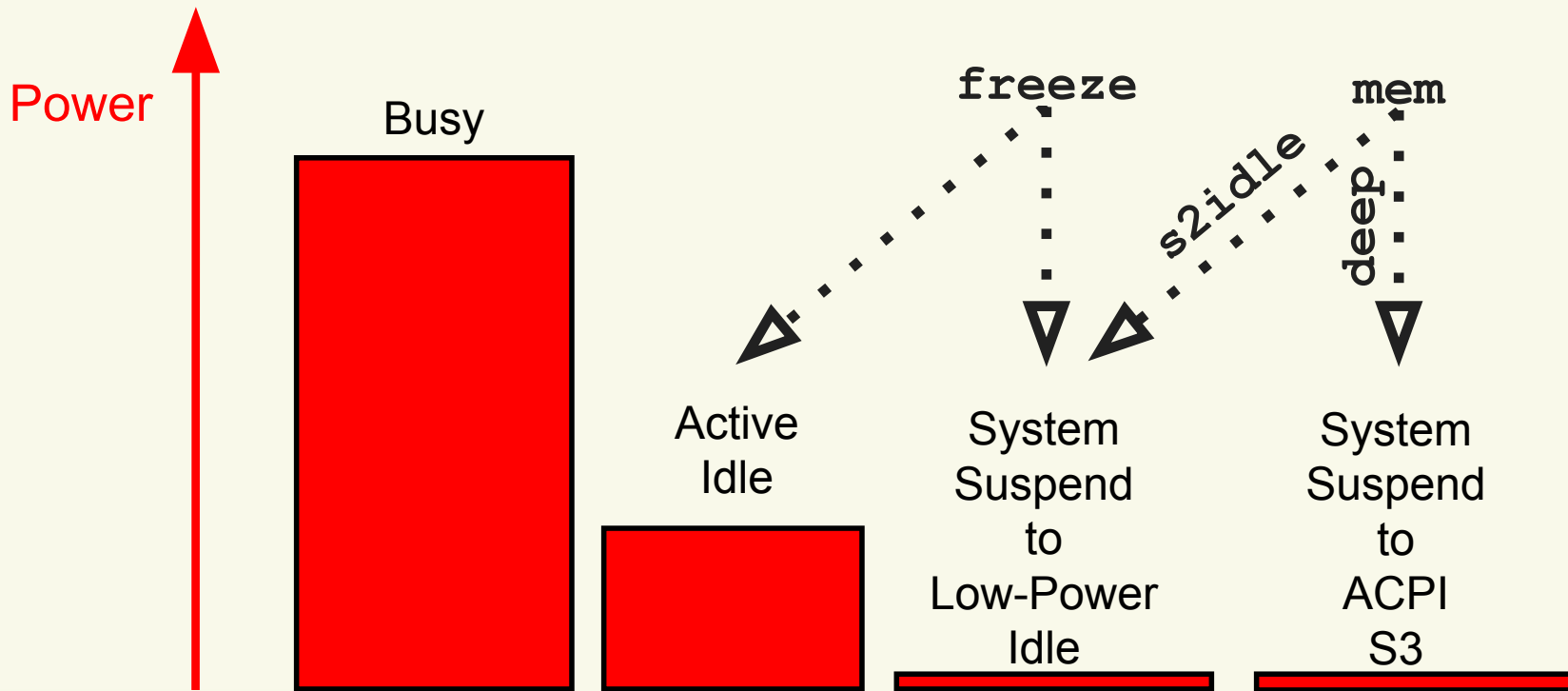


Linux System Suspend to "mem" Default

```
$ cat /sys/power/mem_sleep  
s2idle [deep]
```



System Suspend Options



Part 2 - Sleepgraph



Get & Install Sleepgraph

```
$ cd ~/src/linux
```

```
$ cd tools/power/pm-graph
```

```
$ git clone http://github.com/intel/pm-graph
```

```
$ cd pm-graph
```

```
$ sudo make install
```



Run Sleepgraph

```
$ sleepgraph -m freeze -dev -rtcwake 5 -o outdir
```

Documentation

<https://www.intel.com/content/www/us/en/developer/topic-technology/open/pm-graph/overview.html>

Usage

<https://www.intel.com/content/www/us/en/developer/articles/technical/usage.html>

README

<https://github.com/intel/pm-graph/blob/master/README>



Run Sleepgraph - Screenshot

```
lenb@lenb-Dell-XPS-13-9315:~/lpc23$ sudo sleepgraph -m freeze -dev -rtcwake 5 -o outdir
Checking this system (lenb-Dell-XPS-13-9315)...
  have root access: YES
  is sysfs mounted: YES
  is "freeze" a valid power mode: YES
  is ftrace supported: YES
  are kprobes supported: YES
  timeline data source: FTRACE (all trace events found)
  is rtcwake supported: YES
  optional commands this tool may use for info:
    turbostat: FOUND
    mcelog: FOUND
    lspci: FOUND
    lsusb: FOUND
    netfix: FOUND
os-version          : Ubuntu 23.04
baseboard-manufacturer : Dell Inc.
baseboard-product-name  : 021965
baseboard-serial-number : /2YZP6M3/CNCMK0022F0035/
baseboard-version      : X03
bios-release-date      : 04/12/2023
```



View Results

```
$ sudo sleepgraph -m freeze -dev -rtcwake 5 -o outdir
```

```
$ cd outdir
```

```
$ google-chrome *.html
```

```
$ chromium-browser *.html
```

```
$ firefox *.html
```



SleepGraph v5.11

lenb-Dell-XPS-13-9315 6.5.0 freeze October 07 2023, 01:18:06 PM

Dell Inc. XPS 9315 with 12th Gen Intel(R) Core(TM) i7-1260U

Kernel Suspend Time: **517.553 ms**

freeze time: **3822 ms**

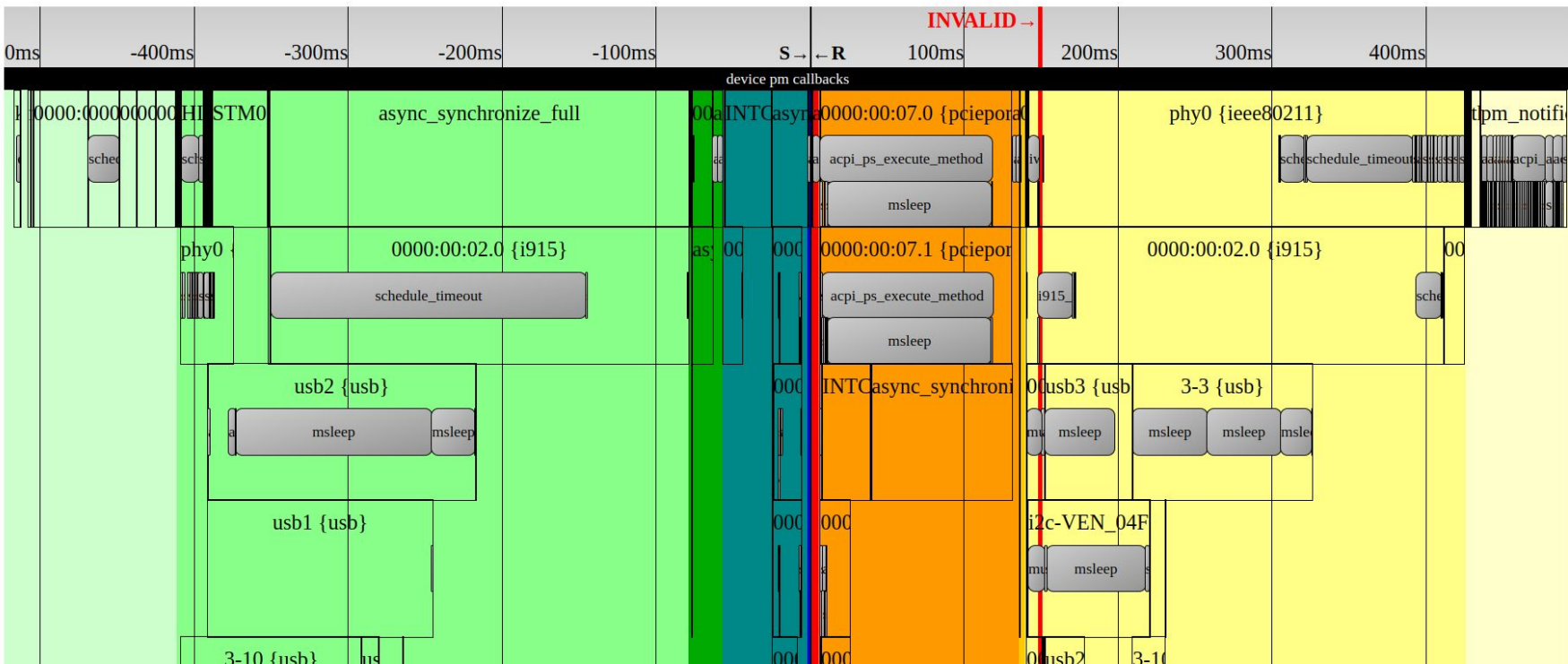
Kernel Resume Time: **489.998 ms**

Device Detail

ZOOM IN +

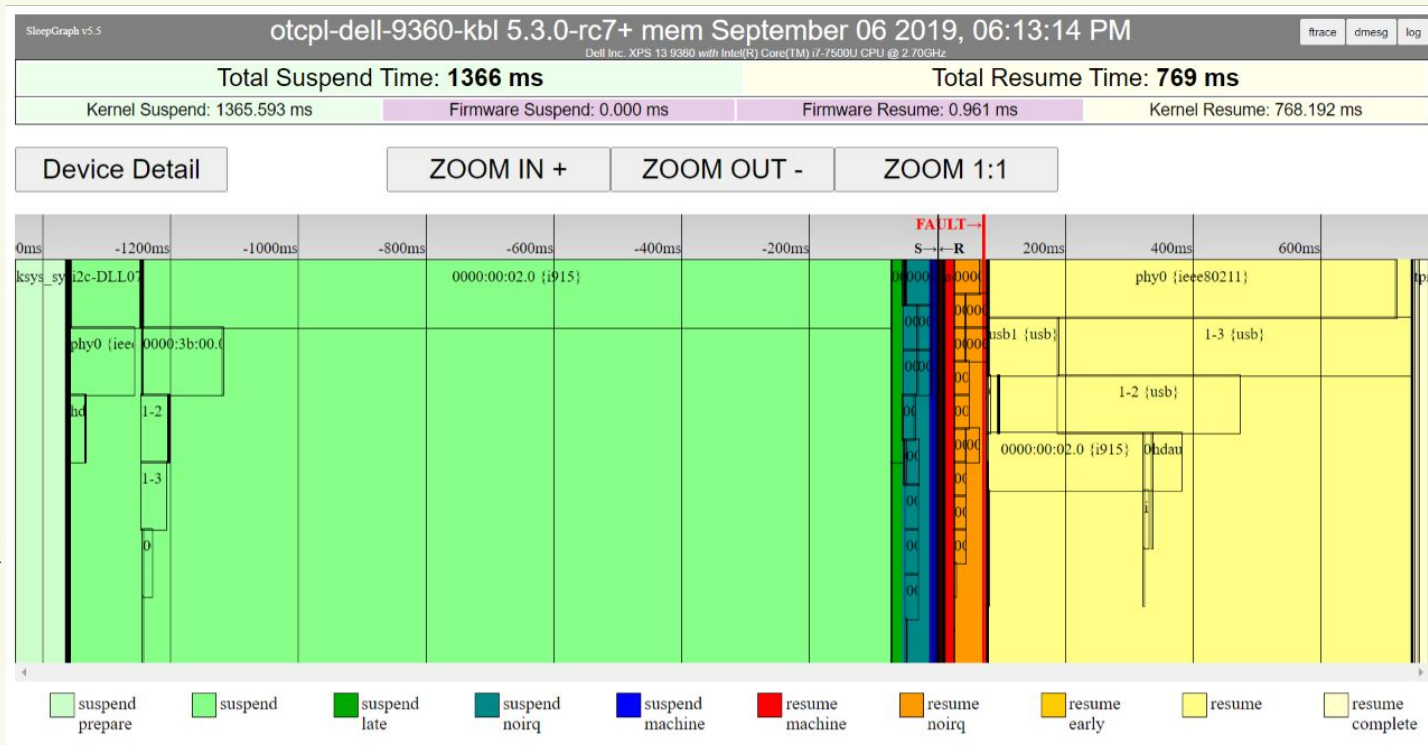
ZOOM OUT -

ZOOM 1:1



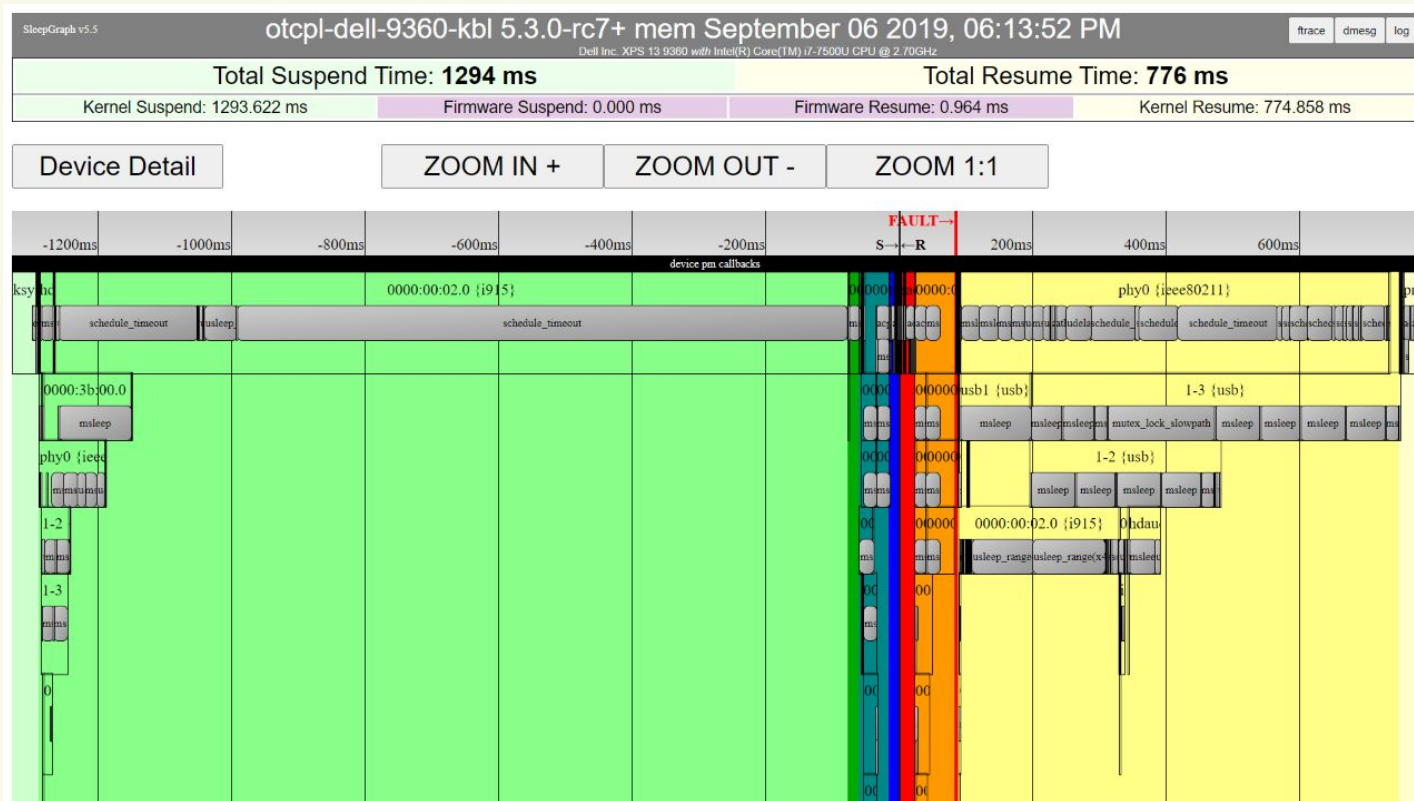
Basic html timeline shows driver callbacks, trace marks

Blocks are driver callbacks during each suspend phase. Trace marks show common calls such as ksys_sync and CPUon/off. Click them to see more detailed info.



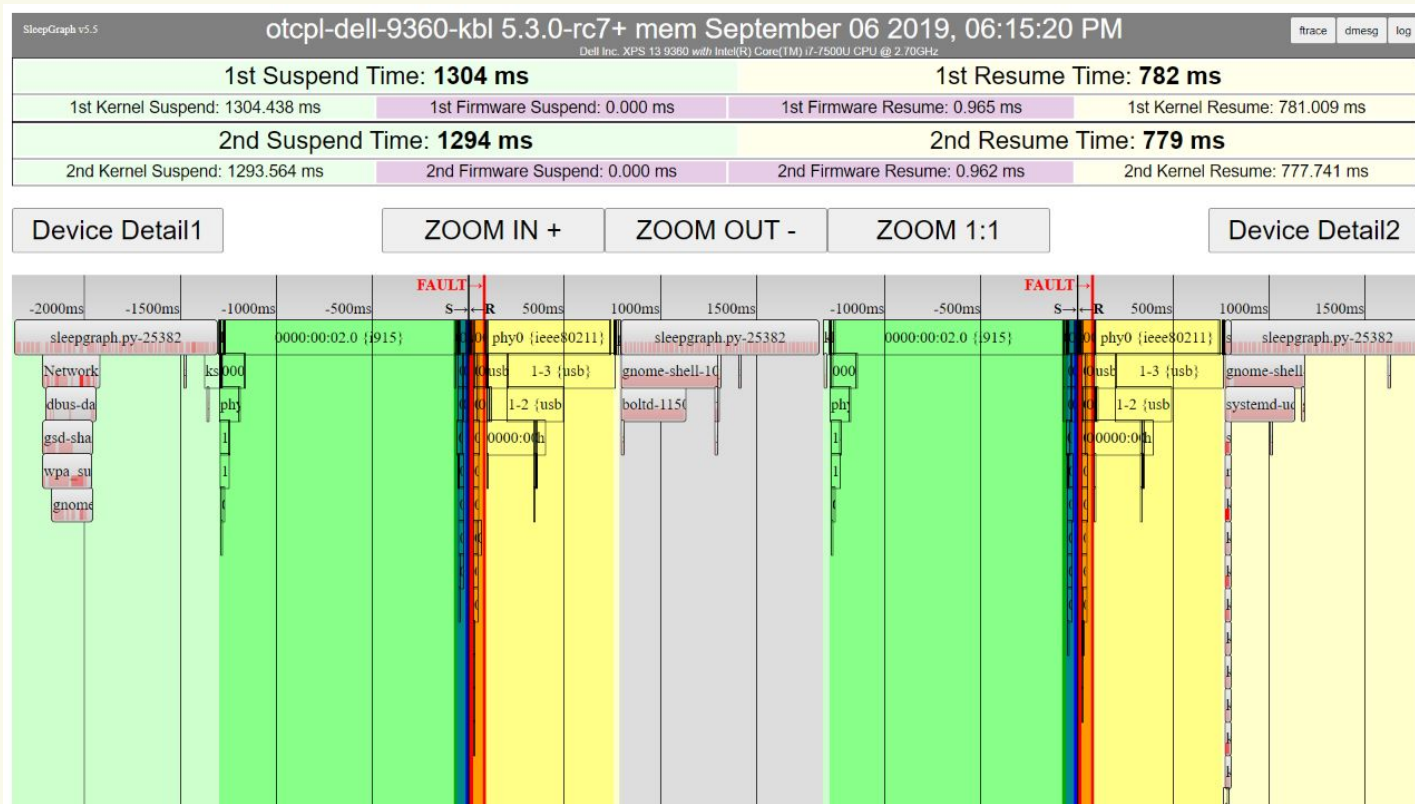
Using -dev mode to visualize function calls with kprobes

Gray bars show function calls, hover to see function args, return values, and parent caller. Enabled with “-dev”.



Using -proc mode to visualize user processes

Rounded bars show processes with CPU usage in red. Enabled with the “-proc” arg. “-x2” allows two back to back suspends.



Run & View Sleepgraph Stress Test Results

```
$ sleepgraph -dev -m freeze -rtcwake 5 -o stress.out  
-multi 3 0 -wifi
```

```
$ cd stress.out
```

```
$ google-chrome summary.html
```

```
$ chromium-browser summary.html
```

```
$ firefox summary.html
```



```

lenb@lenb-Dell-XPS-13-9315:~/lpc23$ sudo sleepgraph -dev -m freeze -rtcwake 5 -o stress.out -multi 3 0 -wifi
Checking this system (lenb-Dell-XPS-13-9315)...
  have root access: YES
  is sysfs mounted: YES
  is "freeze" a valid power mode: YES
  is ftrace supported: YES
  are kprobes supported: YES
  timeline data source: FTRACE (all trace events found)
  is rtcwake supported: YES
  optional commands this tool may use for info:
    turbostat: FOUND
    mcelog: FOUND
    lspci: FOUND
    lsusb: FOUND
    netfix: FOUND
TEST (1/3) START
SUSPEND START
RESUME COMPLETE
TEST (1/3) COMPLETE -- Duration 13.1s
TEST (2/3) START - Avg Duration 13.1s, Time left 0:00:26
SUSPEND START
RESUME COMPLETE
TEST (2/3) COMPLETE -- Duration 13.0s
TEST (3/3) START - Avg Duration 13.1s, Time left 0:00:13
SUSPEND START
RESUME COMPLETE
TEST (3/3) COMPLETE -- Duration 12.9s
Generating a summary of folder:
/home/lenb/lpc23/stress.out
Summary files:
  summary.html          - tabular list of test data found
  summary-devices.html  - kernel device list sorted by total execution time
  summary-issues.html   - kernel issues found sorted by frequency
lenb@lenb-Dell-XPS-13-9315:~/lpc23$ █

```



sleepgraph -multi: summary.html

lenb-Dell-XPS-13-9315 6.5.0 freeze 3 (3 tests: 3 pass)																
#	Mode	Host	Kernel	Test Time	Result	Issues	Suspend	Resume	Worst Suspend Device	SD Time	Worst Resume Device	RD Time	PkgPC10	SysLPI	Wifi	Detail
3	FREEZE	Suspend Avg=478.939 Min=465.786 Med=477.435 Max=493.597 Resume Avg=505.485 Min=501.072 Med=504.443 Max=510.939														
1	freeze	lenb-Dell-XPS-13-9315	6.5.0	2023/10/07 13:24:35	pass	INVALID S2WAKEx2	493.597 ms	504.443 ms	0000:00:02.0 {i915} (async)	291.606 ms	0000:00:02.0 {i915} (async)	292.731 ms	27.65%	20.66%	wlp0s20f3:iwlfwif:8086:51F0 reconnected 0.19	html
2	freeze	lenb-Dell-XPS-13-9315	6.5.0	2023/10/07 13:24:48	pass	INVALID	477.435 ms	510.939 ms	0000:00:02.0 {i915} (async)	288.625 ms	0000:00:02.0 {i915} (async)	288.074 ms	28.51%	20.25%	wlp0s20f3:iwlfwif:8086:51F0 reconnected 0.30	html
3	freeze	lenb-Dell-XPS-13-9315	6.5.0	2023/10/07 13:25:01	pass	INVALID	465.786 ms	501.072 ms	0000:00:02.0 {i915} (async)	279.016 ms	0000:00:02.0 {i915} (async)	291.770 ms	28.10%	20.93%	wlp0s20f3:iwlfwif:8086:51F0 reconnected 0.20	html



Output: `sudo ./sleepgraph -multi 10 0`

|— `summary-devices.html`

|— **`summary.html`**

|— `summary-issues.html`

|— `suspend-220906-094933`

| |— `lenb-Dell-XPS-13-9360_freeze_dmesg.txt.gz`

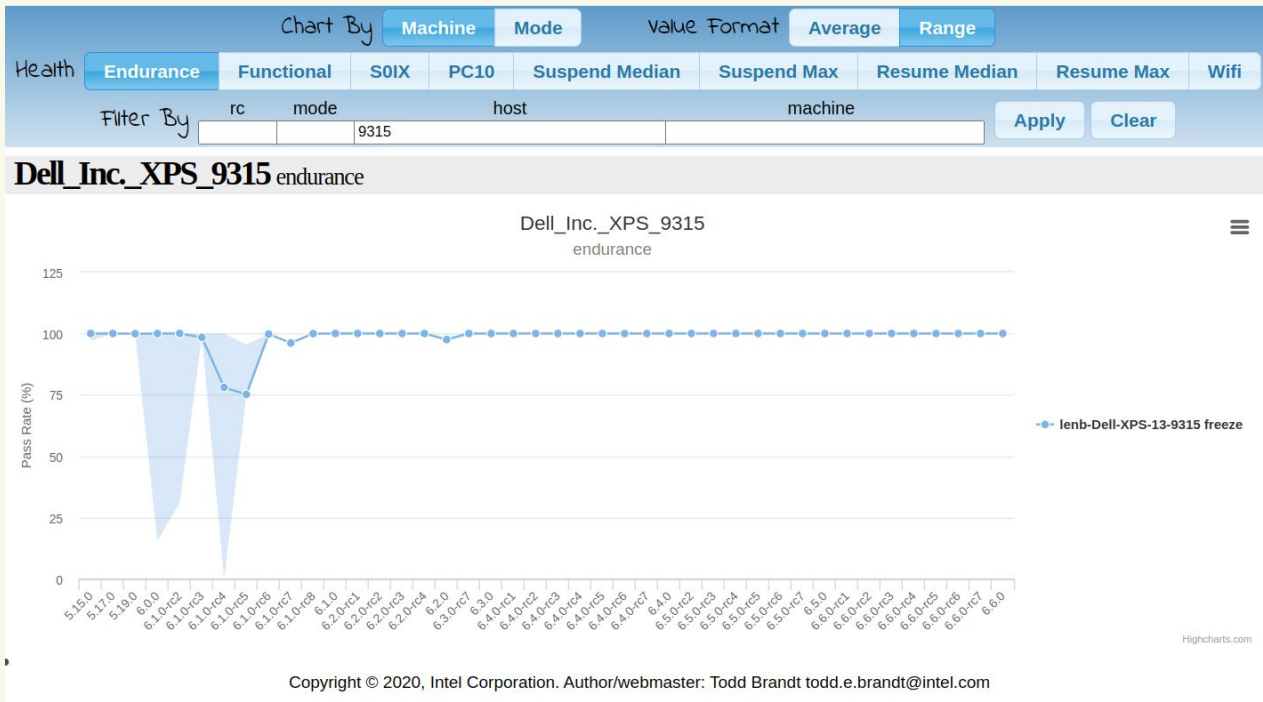
| |— `lenb-Dell-XPS-13-9360_freeze_ftrace.txt.gz`

| |— **`lenb-Dell-XPS-13-9360_freeze.html`**



Trend Charts

Pm-graph repo supports a “stressreport” tool that lets us create datasets and spreadsheets for charting purposes. We have a web UI that employs html graphing tools to visualize kernel performance history.



Endurance Testing Gotchas

System PM config

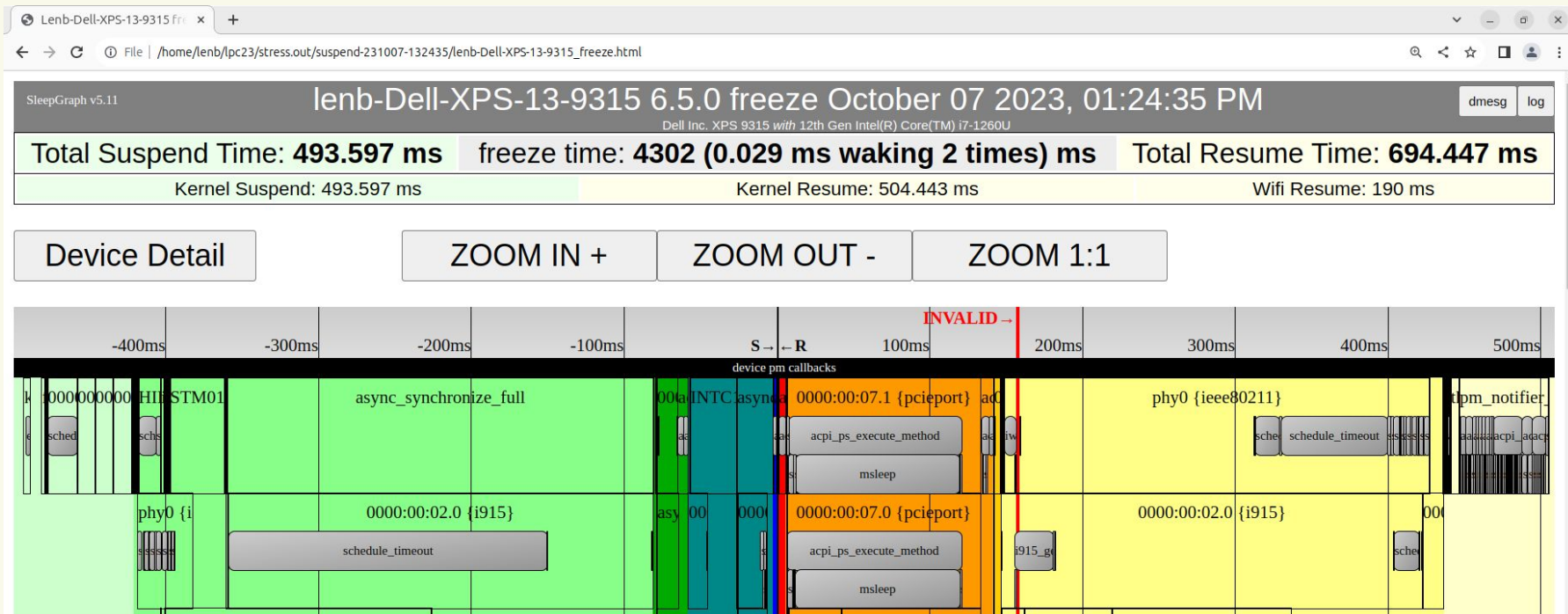
auto suspend, auto screen blank, fstrim.timer strikes at midnight

Network can fail, and stay failed

-netfix added to repair ethernet/wifi after testing through NetworkManager



sleepgraph -wifi



Part 3 - Learnings / Discussion / Questions

Goals Evolve...

job #1: must be 100% reliable -- users must trust it

job #2: must be fast enough to be attractive to users

Today:

1-second laptop resume is considered "good enough"

touch screen needs "instant on/off" perception

Dark resume requires low tax for suspend+resume

Future: ?



How much Endurance is "Enough"?

<u>Iterations</u>	<u>Duration</u>	<u>Issues Found</u>
1	20 sec	Gross functional issues
10	3 min	Most common functional issues
1,000	8 hr	Device transient issues
3,000	24 hr	Typical Endurance Test
10,000	3.5 days	Product Quality



Linux Suspend Architecture: Weakness #1

device drivers can prevent suspend -- but they generally should NOT

(e1000e -fixed)

suspend driver callbacks error return will PREVENT suspend

Generally better to not terminate a suspend in progress



Linux Suspend Architecture: Weakness #2

Run time Device Suspend can make resume FASTER

If resumed system keeps device suspended (eg. display)

but...

Run time Device Suspend can make suspend SLOWER

If run-time resume required before system suspend



Android Uconf Session on Monday:



Linux
Plumbers
Conference | Richmond, VA | Nov. 13-15, 2023

Improving suspend/resume

Saravana Kannan <saravanak@google.com>



Endurance Lab Lessons

consistency is difficult to maintain (WIFI router, BIOS changes, device FW, distro updates etc. -- TEMPERATURE, time of day!)

Long endurance tests are beneficial, but more more machines are more beneficial

In lab: remote/automatic power reset is invaluable

USB ethernet is a PITA -- wifi more reliable

Not all WIFI is created equal -- repeated instability with some chips

wired ethernet: 1 device could prevent suspend (driver fixed)

performance regressions happen: console_suspend/printk, SATA...

Majority of issues are device specific. need a WIDE net to test all devices



Recent kernel bugs caught & fixed by stress testing

[Bug 217955](#) - Performance regression: resume_console takes 100ms longer in S2idle/S3 resume in v6.6-rc1

- Discovered bump in resume time by charting median resume on all machines

[Bug 217804](#) - REGRESSION WITH BISECT: TPM patch breaks S3 on some Intel systems

- Discovered AMD specific TPM commit that interfered with S3 on Intel systems

[Bug 216216](#) - [BISECT INCLUDED] PNP0501 serial driver takes almost 2 seconds to suspend/resume

- Discovered bump in suspend time by charting median suspend on all machines

[Bug 216600](#) - RTC wakealarm file is missing in v6.1-rc1

- Discovered wakealarm file missing which sleepgraph uses to wake from suspend.

[Bug 208733](#) - s2idle freeze wakes from timekeeping_freeze at least 100 times with ec_no_wakeup enabled

- Discovered ec_no_wakeup kernel arg ceased to work correctly in S2idle



Hibernate to Disk - Future Revival?

1st reliable system suspend method

Stable, but disabled by default in Fedora, Ubuntu, configurable in SuSE

default (bare metal) images are not encrypted

Renewed Interest?

Migration use-case on VMM

filling crypto gap may lead to broader enablement



Future Trends

Endurance achieved, must assure it doesn't regress

Dark suspend/resume latency remains particularly important

On Intel, Low-Power Idle residency is mandatory

ACPI S3 use dwindling

Hibernate of interest again?



What should happen in the next 10 years?

- Legacy is still in production: hibernate and S3
- Must continuously regression test (function, power and performance)
- Must test on more different systems and peripherals
 - (Drivers for new devices must support suspend/resume)

