

So you're a Linux kernel developer?
Name all subsystems.

January 23, 2021



OSTBAYERISCHE
TECHNISCHE HOCHSCHULE
REGENSBURG

Pia Eichinger

- ▶ Student at the University of Applied Sciences in Regensburg (OTH)
- ▶ Email: pia.eichinger@st.oth-regensburg.de
- ▶ Research interests: organisational and maintenance structure of the Linux kernel





Ralf Ramsauer

- ▶ PhD student
- ▶ ralf.ramsauer@othr.de



Stefanie Scherzinger

- ▶ Professor for scalable database systems
- ▶ stefanie.scherzinger@uni-passau.de



Wolfgang Mauerer

- ▶ Professor for theoretical computer science
- ▶ Siemens AG, Corporate Research and Technology
- ▶ wolfgang.mauerer@othr.de

Motivation and Goals

- ▶ Formalise/Assess the Linux kernel development process
- ▶ Provide tools to support the open source community
- ▶ Assist/Support Safety Critical Developments

Collaboration Partners

- ▶ ELISA (Linux Foundation)
- ▶ University of Passau

Safety Certification

- ▶ Highly sensitive environment
- ▶ Working software could ensure survival
- ▶ Need to ensure software quality **by all means necessary**

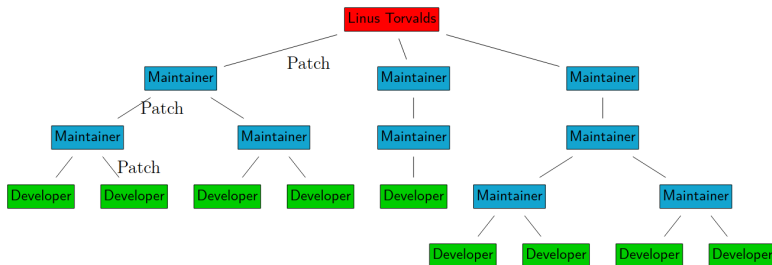
Process and Quality

- ▶ Assumption: Development process impacts quality
- ▶ Process adherence necessary (Degree?)

Linux + Safety Critical Environments

= Major certification problem: Open development process

- ▶ Ex post facto analysis
- ▶ Characterise the process with hindsight (patch integration)
- ▶ Statistical Methods to research/understand the development process



Original Projects

- ▶ Conforming Integration of Patches
- ▶ Patch traversal through maintainers hierarchy

Only problem is ...

- ▶ ... where can I find the maintainers hierarchy?
- ▶ ... where can I find the documentation on all subsystems?
- ▶ ... what exactly is a subsystem even?

Docs: Early-stage Planning

“Again, the MAINTAINERS file is the place to start. [...] not all **subsystems** are represented there.”

ALLWINNER VPU DRIVER

```
M:      Maxime Ripard <mripard@kernel.org>
M:      Paul Kocialkowski <paul.kocialkowski@bootlin.com>
L:      linux-media@vger.kernel.org
S:      Maintained
F:      drivers/staging/media/sunxi/cedrus/
```

ALPHA PORT

```
M:      Richard Henderson <rth@twiddle.net>
M:      Ivan Kokshaysky <ink@jurassic.park.msu.ru>
M:      Matt Turner <mattst88@gmail.com>
L:      linux-alpha@vger.kernel.org
S:      Odd Fixes
F:      arch/alpha/
```

ALPS PS/2 TOUCHPAD DRIVER

```
R:      Pali Rohár <pali@kernel.org>
F:      drivers/input/mouse/alps.*
```

```
This section contains usage information about media subsystem and  
its supported drivers.
```

```
Please see:
```

```
- :doc:`admin-guide/media/index`  
    for usage information about media subsystem and supported drivers;
```

But...

- ▶ No “Media”/“Media Subsystem” entry in MAINTAINERS
- ▶ Over 100 “subsystems” containing the word “Media”

⇒ Which one?

Definition: Subsystem

- ▶ Entries in MAINTAINERS: **sections**
- ▶ Sections share files (measured in Lines of Code): **thematically related**
- ▶ Grouping of thematically strong related sections: **subsystem**.

No clear listing of subsystems

⇒ Let's find out!

New Projects

- ▶ Conforming Integration of Patches
- ▶ Subsystem Detection

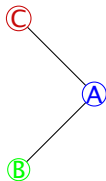
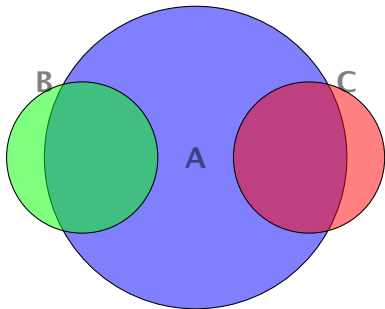
Goal: Find and analyse subsystems

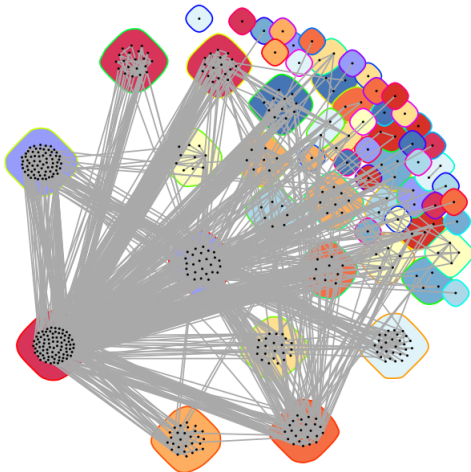
Visualise MAINTAINERS sections and thematical relations

Definition: Section Graph

- ▶ Undirected Graph
- ▶ Vertices: Sections of MAINTAINERS
- ▶ Edge: Do two sections share LoC? Yes $\Rightarrow$ Edge

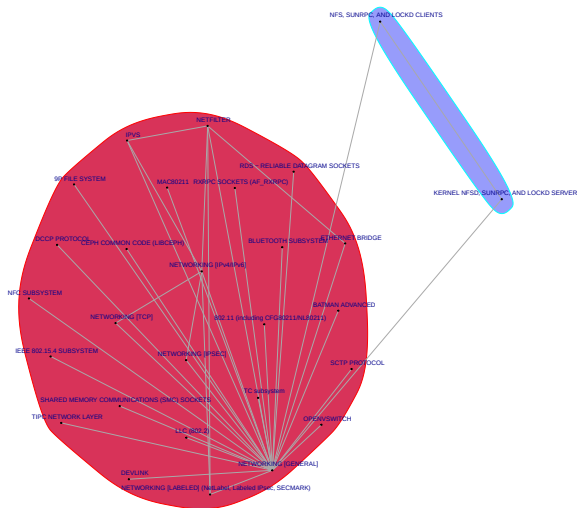
$\Rightarrow$ Detect clusters (subsystems)





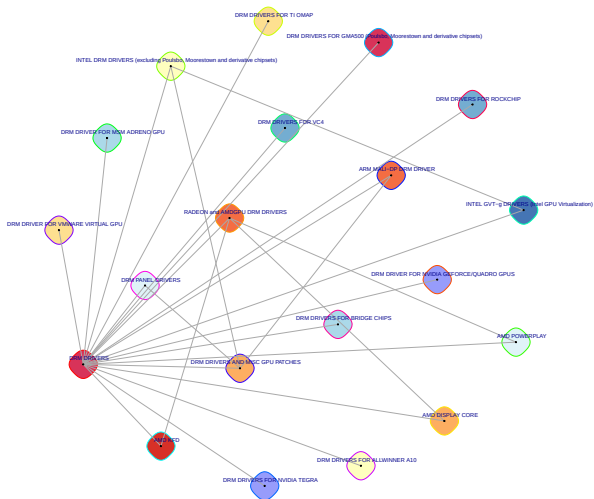
Cluster Discussion

- ▶ Major clusters: **isolate as own graph**
- ▶ “Recluster” again from within: **substructure**



So you're a Linux kernel developer? Name all subsystems.





Conforming Patch Integration

Patch integrated by relevant maintainer (from MAINTAINERS) $\Rightarrow$

Conforming Integration

tl;dr

- ▶ Analyse recent patch integrations
- ▶ Determine conforming patch integration

$\Rightarrow$ Analyse reasons for unconforming patch integration



Important Note

- ▶ Goal: Characterise/Improve process and patch integration
 1. Extract current status
 2. Analyse/Characterise/Discuss
 3. Enhance
- ▶ We do not mean to point fingers!

Results

- ▶ Definition for subsystems
- ▶ Visual representation of MAINTAINERS
- ▶ Fully-automated and sensible subsystem detection
- ▶ Strong argumentation basis for certifications

Future Work

- ▶ Combination: Subsystems and Conforming Patch Integration
- ▶ Cluster discussion on full section graph
- ▶ Visualise earlier versions and subsystem development
- ▶ Easier overview for newcomers (interactive graph)
- ▶ Apply section graph to other open source projects (QEMU, U-Boot, buildroot)
- ▶ Maintainer Graph

All work integrated in <https://github.com/lfd/PaStA>

Thank you for listening!

`pia.eichinger@st.oth-regensburg.de`

`https://github.com/lfd/PaStA`