

# Unifying and improving test regression reporting and tracking

**Gustavo Padovan**  
**Ricardo Cañuelo**



COLLABORA

Open First

# Where we're at

- Multiple kernel-oriented CI systems: KernelCI, CKI, 0-day, syzbot, LKFT
- Common place to aggregate the results: KCIDB
- A de-facto tool to track regressions: Regzbot



# Problems

- Many developers may not be aware of regressions that could be relevant to them
- In contrast, some maintainers get way too many regression reports they're not interested in
- Automatic regression reports may not always point to real regressions
- Current CI systems provide only the most basic info about a regression
- No concept of a regression life cycle in CI systems



# Solution: Improve the quality of regression data provided by CI systems

- Initial data: Failed test, test environment, software version that caused the failure, logs

# Improve the quality of regression data provided by CI systems

- Post-processing this data we can generate more elaborate info
  - Is the test failing intermittently? Unreliable test -> possible false positive
  - Were other tests running in the same lab/setup also failing around the same time? Possible infrastructure error -> not a regression
  - Results of the same test on the same kernel version but with different kernel configs, on different boards, etc. -> narrow down the possible causes
  - Log analysis: matching and classification of regression types



# Improve the quality of regression data provided by CI systems

- Adding this extra layer of intelligence provides:
  - Automatic filtering of false positives
  - Automatic categorization
  - Initial pre-triaging stage



# Solution: Improve the quality of regression reports

- Report filtering: which reports to send to whom?
- Implement customizable user subscription and muting of reports
- Ideally: web front-end to query reported regressions, check their status and manage report subscriptions



# Solution: model and track a regression life cycle

- Regressions aren't static: they're detected, then reported, then investigated and fixed
- CI systems don't keep track of this
- Communication is currently one-directional from CI to users





# Model and track a regression life cycle

- Missing information from a regression:
  - did anyone claim it?
  - is it still failing?
  - is this linked to other regressions?
  - was this found to be irrelevant / intended / not a regression?
- Regzbot provides these features and is the de-facto regression tracker for the kernel
- Integration with CI efforts needed





**Thank you!**



COLLABORA

**Open First**



**We are hiring**  
**[col.la/careers](https://col.la/careers)**



COLLABORA

Open First