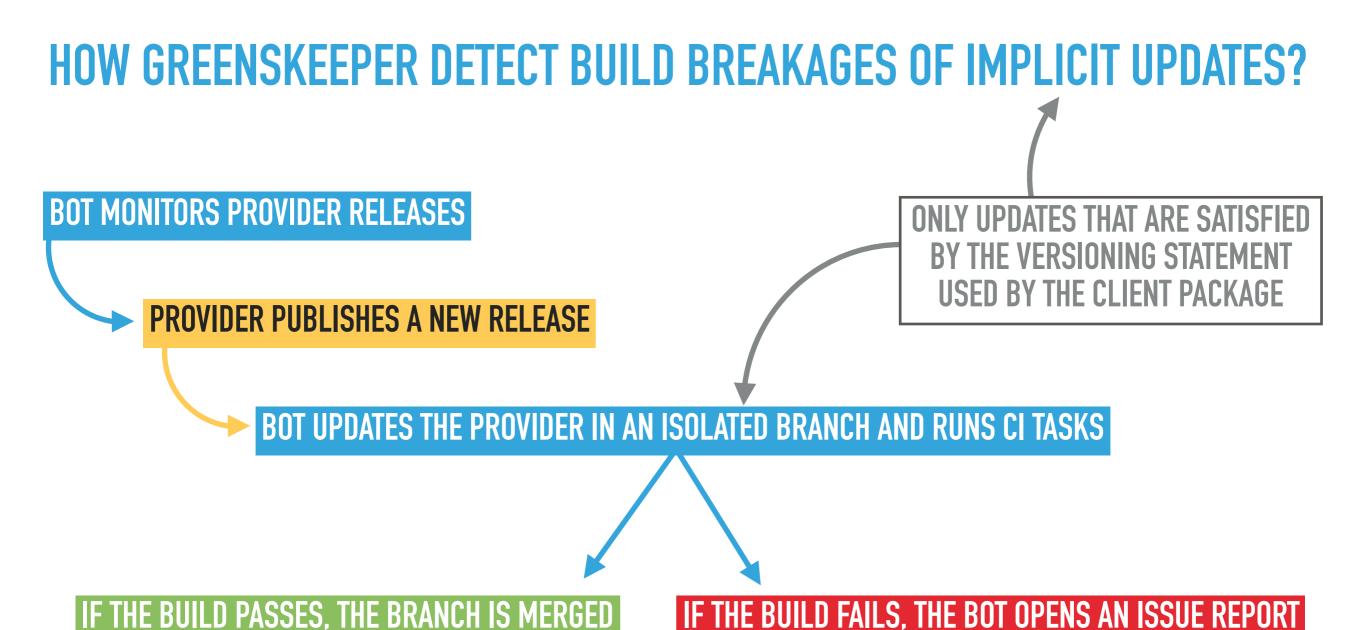
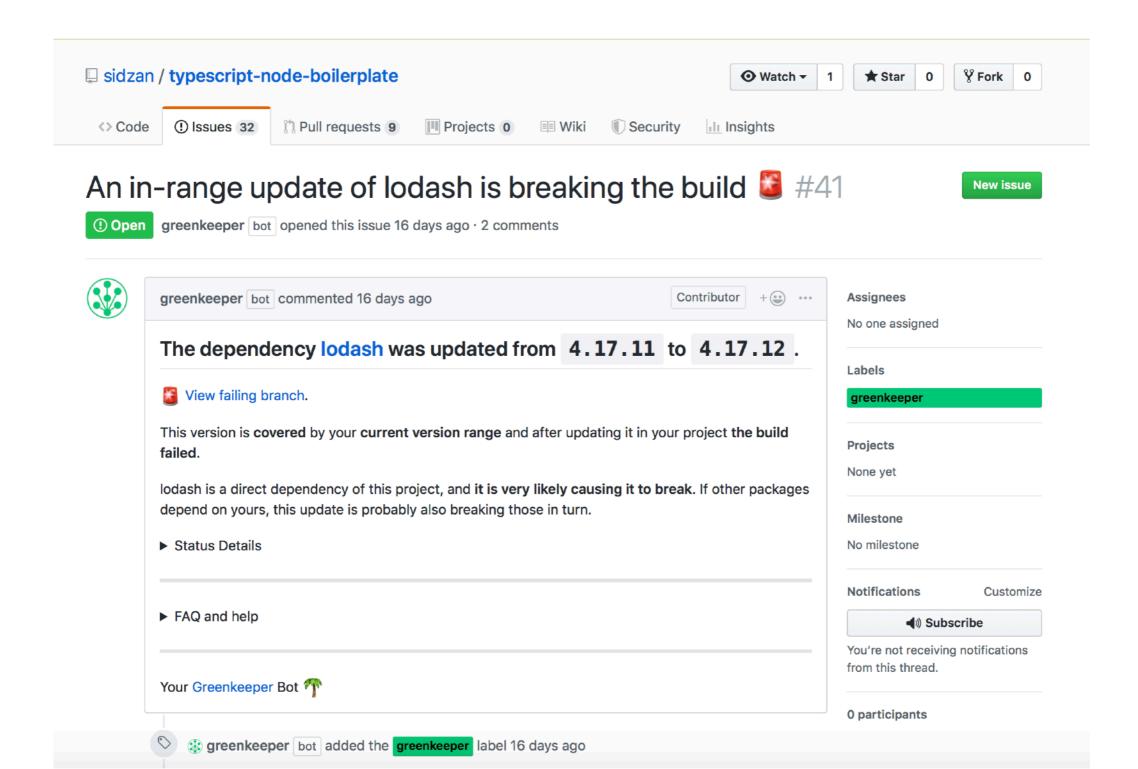
ISSUES OPENED BY GREENKEEPER BOT

DATA SET PRESENTATION

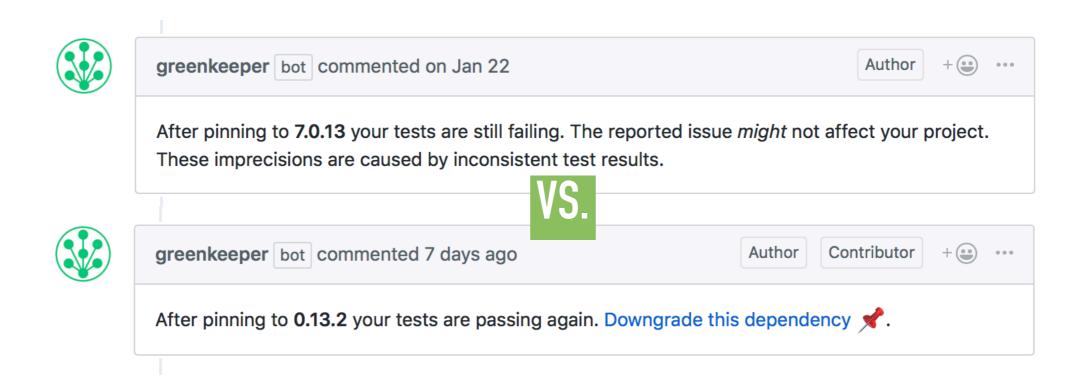


A TYPICAL ISSUE REPORT OPENED BY GREENKEEPER

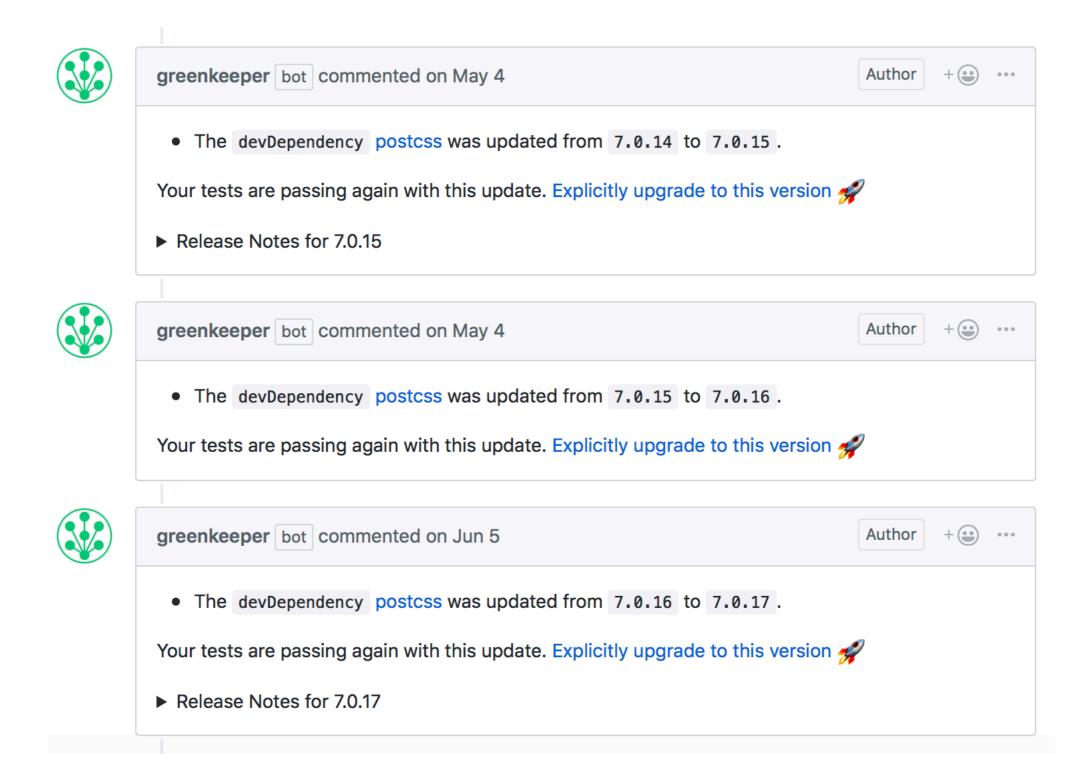


PINNING IS RECOMMENDED WHEN IT FIXES THE BROKEN BUILD

Pinning has an impact on the future updatability of the provider



WHEN A PINNED DEPENDENCY PUBLISHES A RELEASE THE CLIENT PACKAGE IS NOTIFIED



WHICH DATA DO WE HAVE?

WE OBTAINED ALL ISSUE REPORTS THAT HAVE THE STANDARD TITLE

▶ 82,399 issue reports



an in-range update is breaking build in:title

- 152,257 events associated with the issue reports
 - Closed/Reopened (62,154/144)
 - Commits with reference to the issue report (1,388)
 - Assignments to a contributor (1,172)
 - Mention of a GH user (389)
 - Subscriptions for notification (321)
 - Milestoned/Demilestoned (314/51)
 - Among others ...

WE EXTRACTED THE NAME AND VERSION OF THE PROVIDER THAT BROKE THE BUILD

THERE ARE SEVEN VARIATIONS OF ISSUE REPORT BODY. WE USE A DIFFERENT REGULAR EXPRESSION OVER EACH VARIATION

greenkeeper bot commented on Feb 7, 2018

Branch

Type

build failed.

or publishing.

Dependency

Current Version

Version 1.0.2 of grunt was just published.

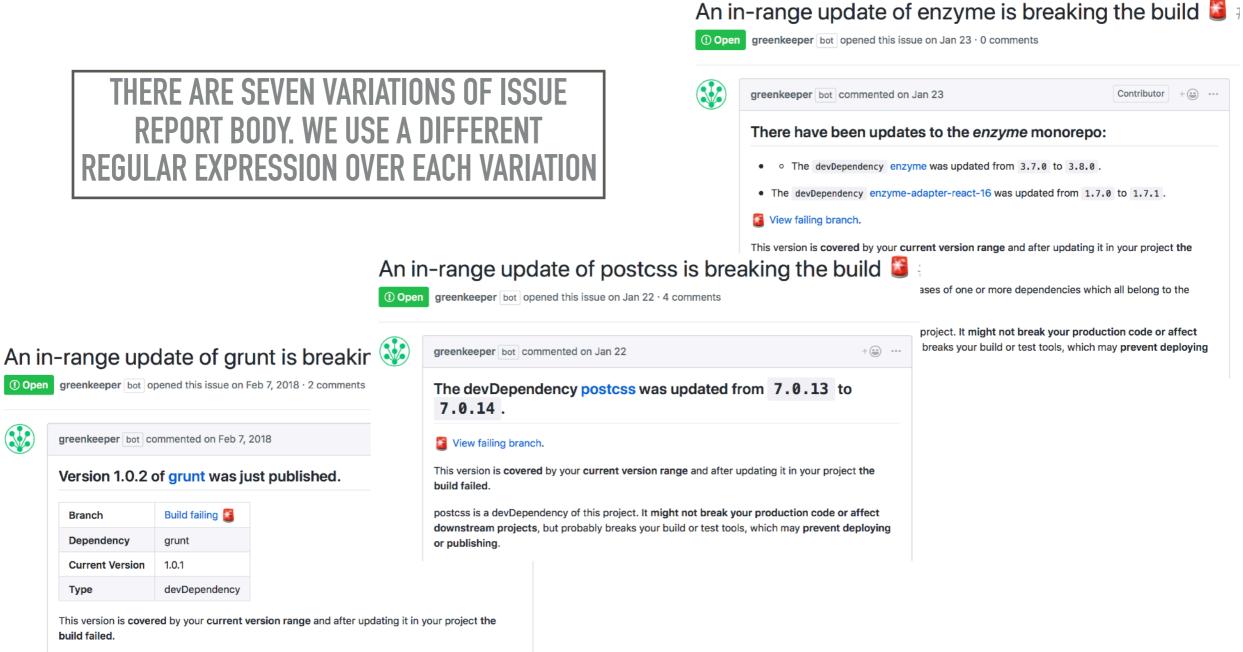
Build failing

devDependency

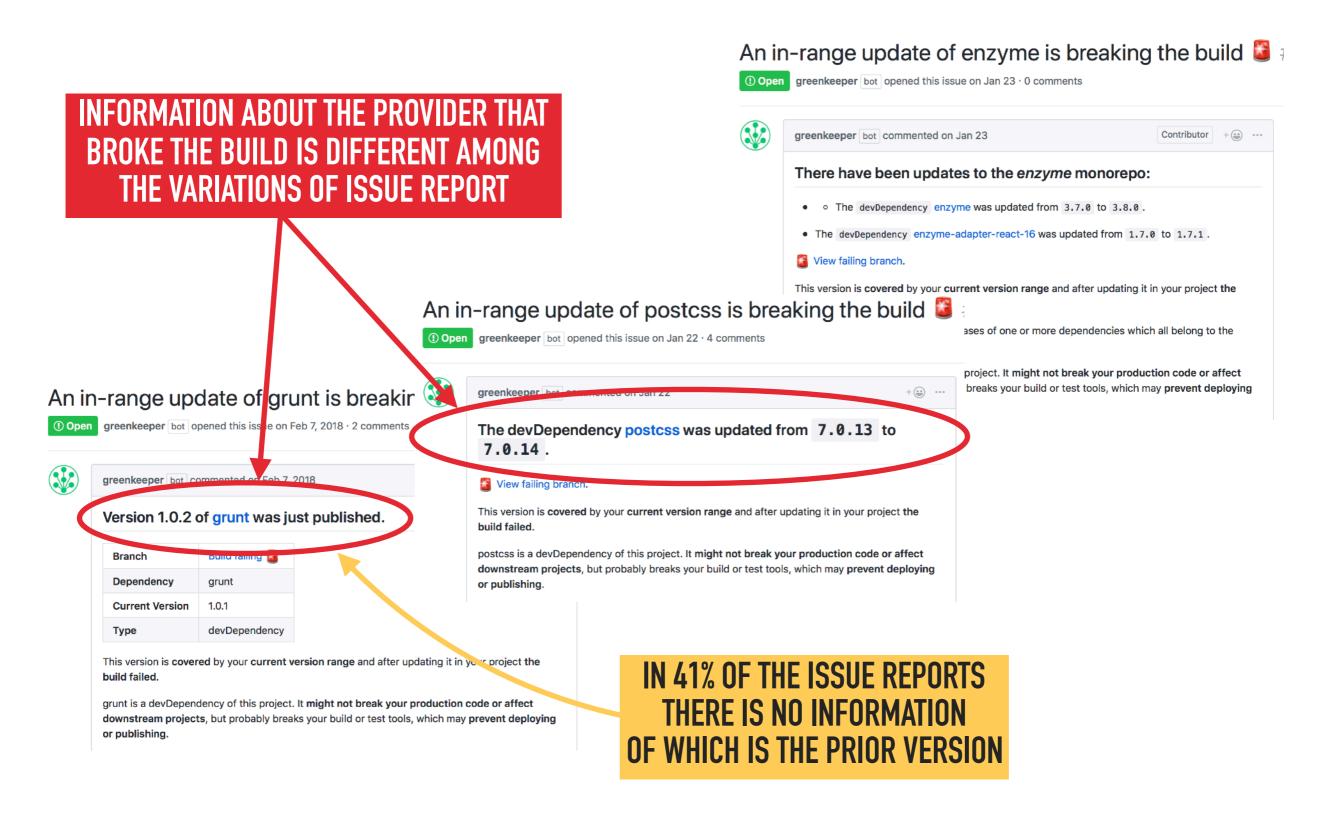
grunt is a devDependency of this project. It might not break your production code or affect downstream projects, but probably breaks your build or test tools, which may prevent deploying

arunt

1.0.1



WE EXTRACTED THE NAME AND VERSION OF THE PROVIDER THAT BROKE THE BUILD



WE OBTAINED ALL SUCH COMMITS

- Commit message
 - Clear vs. Tangled changes
 - "Update dependencies make greenkeeper happy, closes #17, closes #18"
 - "feat(base): Add `function-paren-newline` rule\n\nFix #25"
- Modified files
- Author and committer
- Commit size
 - ▶ Large vs. Regular commits
 - Number of added and deleted lines of code
 - Number of files

CHALLENGES

WE DON'T KNOW IF THE PACKAGE THAT USES THE BOT ON GH IS IN NPM

- There are 10,496 projects using Greenkeeper in our dataset
- If we want to study how the projects (bot users) deploy versions in npm as client packages ...
 - ... then we need to find the correspondent package in npm
- We just have the repository of the project, not it's name in npm
- The package.json file hosted on GH has the name of the package
 - However, life is harder than this!

LINKING THE CLIENT PACKAGE IN GITHUB WITH THE PACKAGE IN NPM IS CHALLENGING

- Revision control on Git and deployment on npm is out-ofsync
 - Package.json file have different states in npm and Git



BEST EFFORT TO FIND WHICH IS THE NPM CLIENT PACKAGE HOSTED ON GITHUB

- Find the latest revision of the package.json file hosted on GH
- Find the package.json file hosted on npm for same package name
- Check the following matches:
 - Repository (44% matches)
 - Author (33% matches)
 - Latest version of client package on GH exists on npm (51.7% match)
 - Name of dependencies (61% matches)

OPPORTUNITIES

HOW CAN I MAKE MY BOT BETTER?

- How much attention is being given to issues opened by the bots?
- How often developers of client packages close the issue without any event (discussion, commit, etc.)?
- How often the bot send false alarms?
- Do developers accept the bot recommendations of pinning/implicit updates?

HOW SHOULD I FIX A BUILD BREAKAGE?

- Do client packages need large changes to cope with the issue opened by the bot?
- How long does a client package takes to react to the issue?
- How long a pinned provider remains without being updated?
- How much people is generally allocated to solve the issues?
- What is going on closed/opened issues with large discussions?

AM I BREAKING MY CLIENTS CODE?

- How often do provider updates break a client package build?
- What types of updates (patch, minor, major) most often break the clients build?
- What types of change do client packages need to perform to recover from a build breakage?