

An Exploratory Study Into the Prevalence of Botched Code Integrations

[@wardmuylaert](#)

[@oniroi](#)

Ward Muylaert and Coen De Roover

Software Languages Lab
Vrije Universiteit Brussel
Brussels, Belgium




VRIJE
UNIVERSITEIT
BRUSSEL

Textual Conflicts

```
function foo(x) {  
  if (x < 0) {  
    return -x;  
  }  
  return x;  
}
```

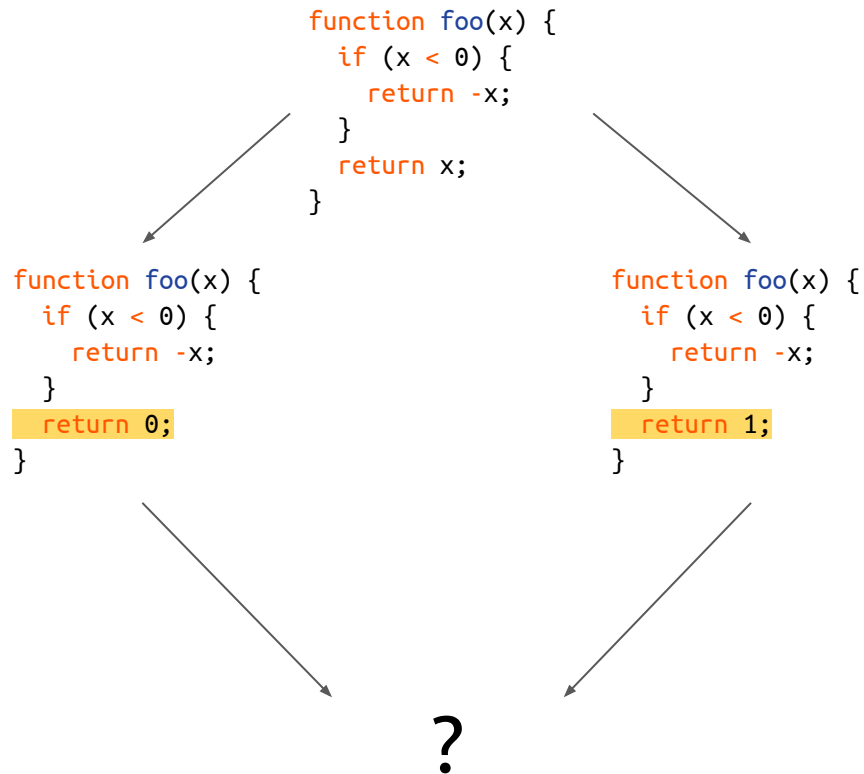
Textual Conflicts

```
function foo(x) {  
  if (x < 0) {  
    return -x;  
  }  
  return x;  
}
```



```
function foo(x) {  
  if (x < 0) {  
    return -x;  
  }  
  return 0;  
}
```

Textual Conflicts

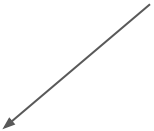


Semantic Conflicts

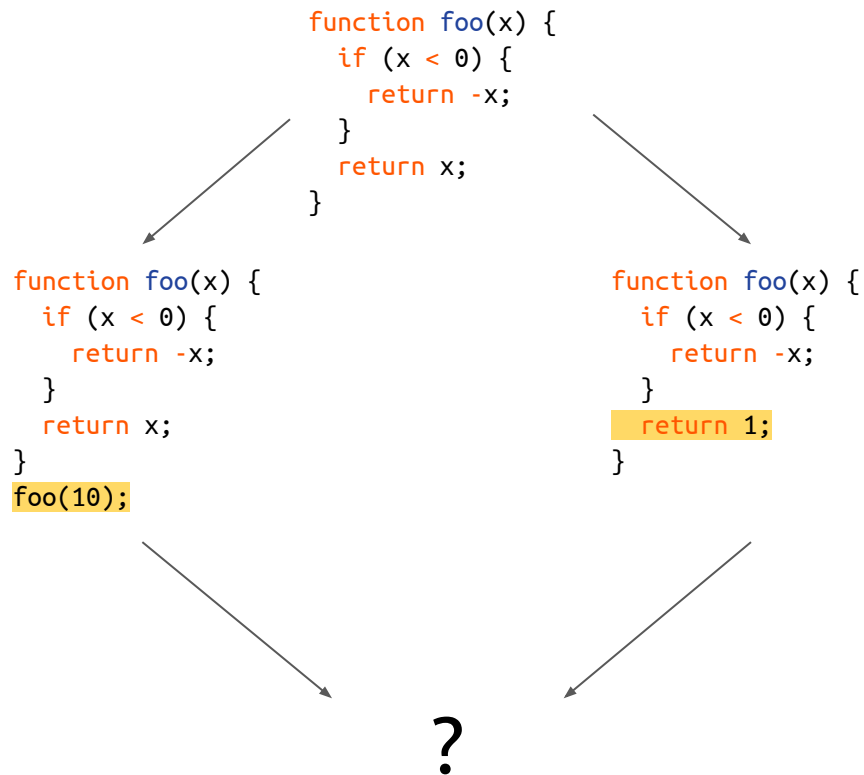
```
function foo(x) {  
  if (x < 0) {  
    return -x;  
  }  
  return x;  
}
```

Semantic Conflicts

```
function foo(x) {  
  if (x < 0) {  
    return -x;  
  }  
  return x;  
}  
  
function foo(x) {  
  if (x < 0) {  
    return -x;  
  }  
  return x;  
}  
foo(-10);
```

A diagram consisting of a single arrow pointing from the argument -10 in the function call foo(-10); to the first function definition of foo(x). This illustrates a semantic conflict where the same function name is used for two different functions, and the resolution is based on the first definition encountered.

Semantic Conflicts



Research Questions

RQ1

How often does code integration lead to semantic conflicts?

RQ2

How much effort is needed to fix semantic conflicts after code integration?

RQ3

How long does it take to fix semantic conflicts after code integration?

GitHub

Commits on Jul 7, 2016



init/Kconfig: keep Expert users menu together ...



Randy Dunlap committed with torvalds 13 hours ago

Commits on Jul 6, 2016



Merge git://git.kernel.org/pub/scm/linux/kernel/git/davem/net ...

torvalds committed 19 hours ago



Merge tag 'sound-4.7-rc7' of git://git.kernel.org/pub/scm/linux/kerne... ...

torvalds committed 20 hours ago

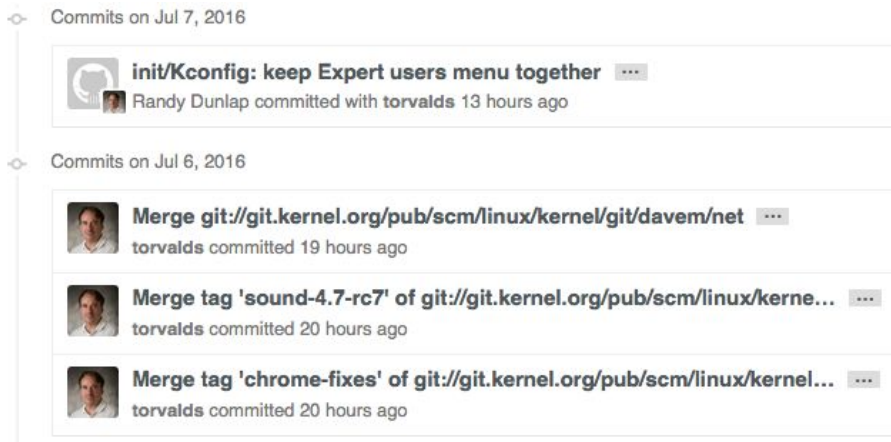


Merge tag 'chrome-fixes' of git://git.kernel.org/pub/scm/linux/kernel... ...

torvalds committed 20 hours ago



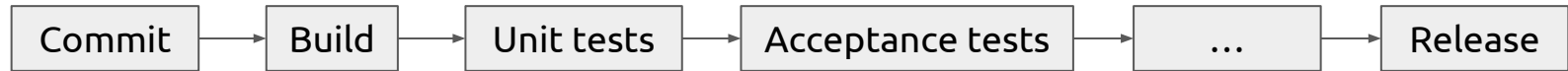
GitHub



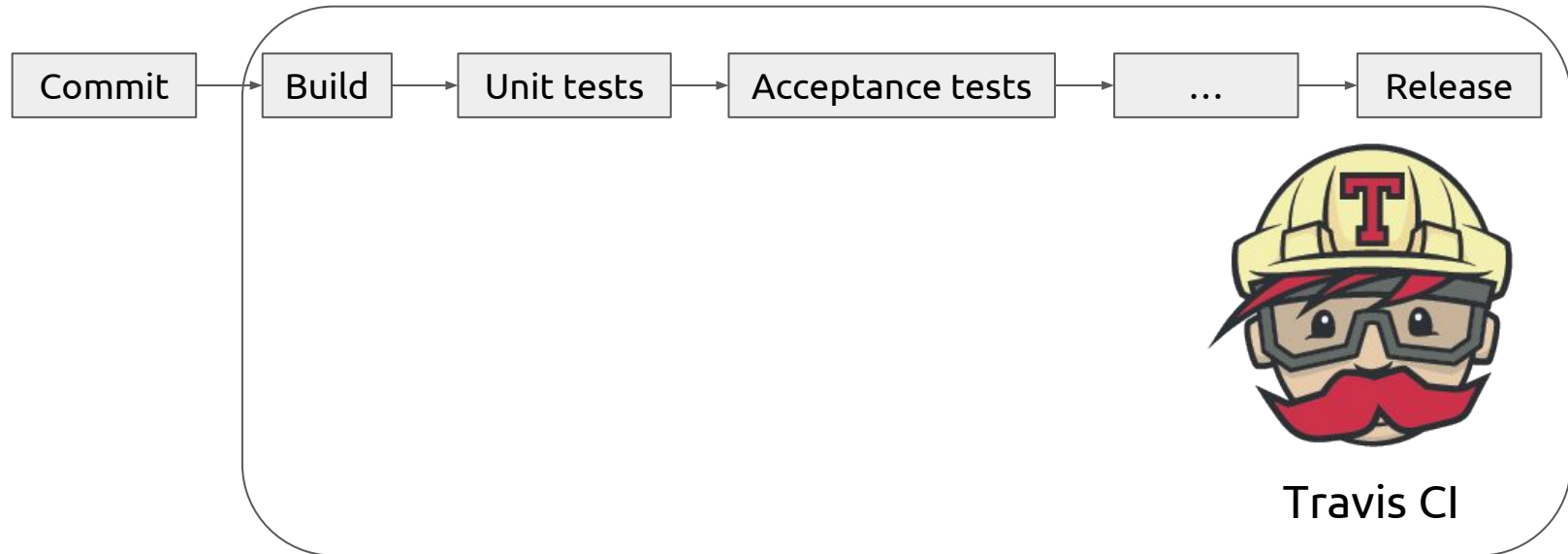
GHTorrent
150GB MySQL dump, 400M commits

Georgios Gousios, The GHTorrent
Dataset and Tool Suite, MSR 2013.

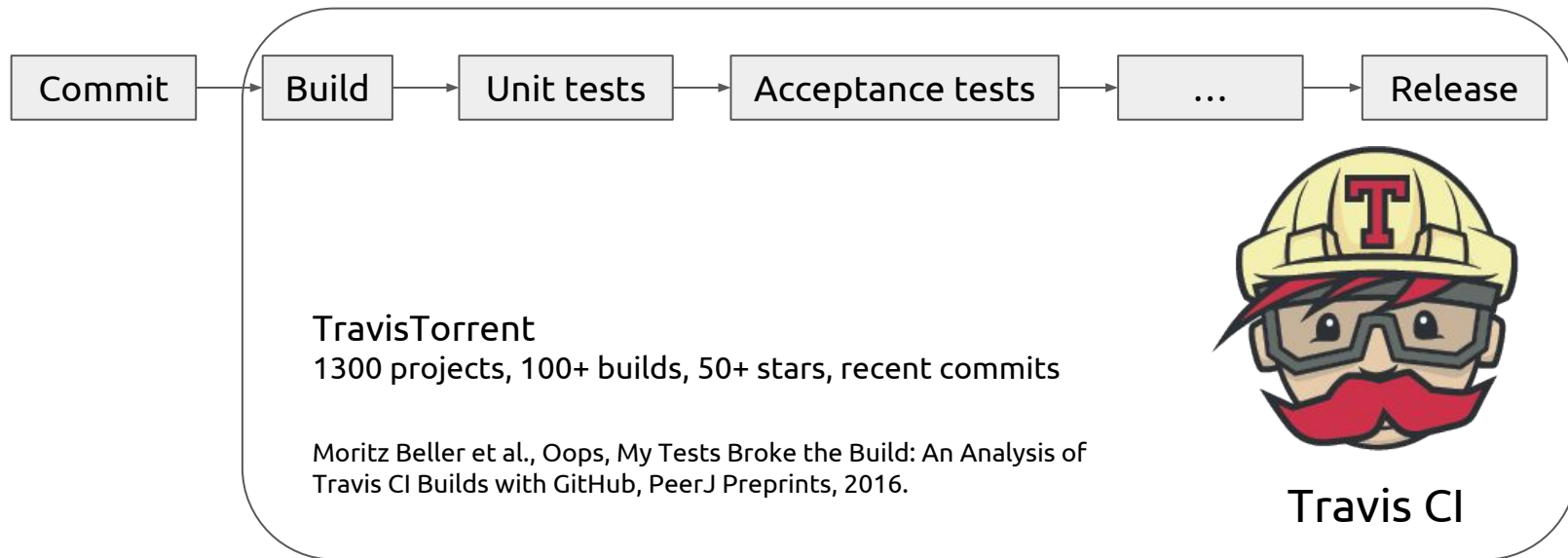
Continuous Delivery



Continuous Delivery



Continuous Delivery



Travis CI API

Builds

state
started_at
finished_at
commit_id
repository_id
...

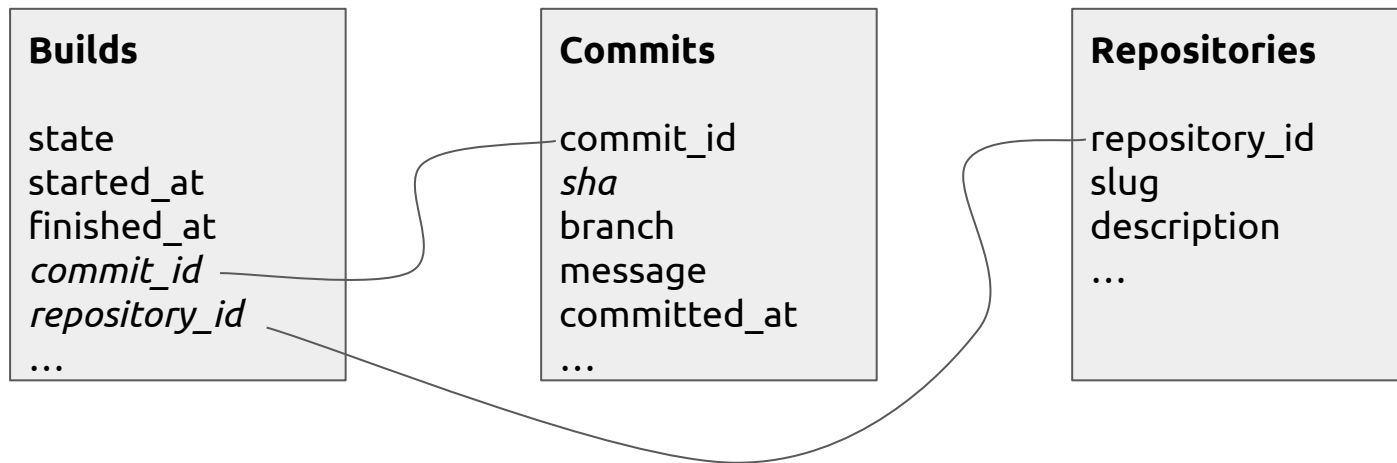
Commits

commit_id
sha
branch
message
committed_at
...

Repositories

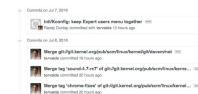
repository_id
slug
description
...

Travis CI API



Travis CI API

GitHub



GoTencent
15GB MySQL dump, 400M commits
Georgios Gionas, The GoTencent
Dataset and Tool Suite, MSR 2013.

Builds

state
started_at
finished_at
commit_id
repository_id
...

Commits

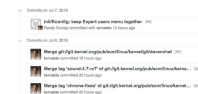
commit_id
sha
branch
message
committed_at
...

Repositories

repository_id
slug
description
...

Travis CI API

GitHub



GitHub
150GB MySQL dump, 400M commits
Georgios Gousios, The GitHub
Dataset and Tool Suite, MSR 2013.

Builds

state
started_at
finished_at
commit_id
repository_id
...

Commits

commit_id
sha
branch
message
committed_at
...

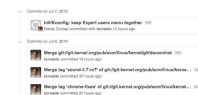
Repositories

repository_id
slug
description
...

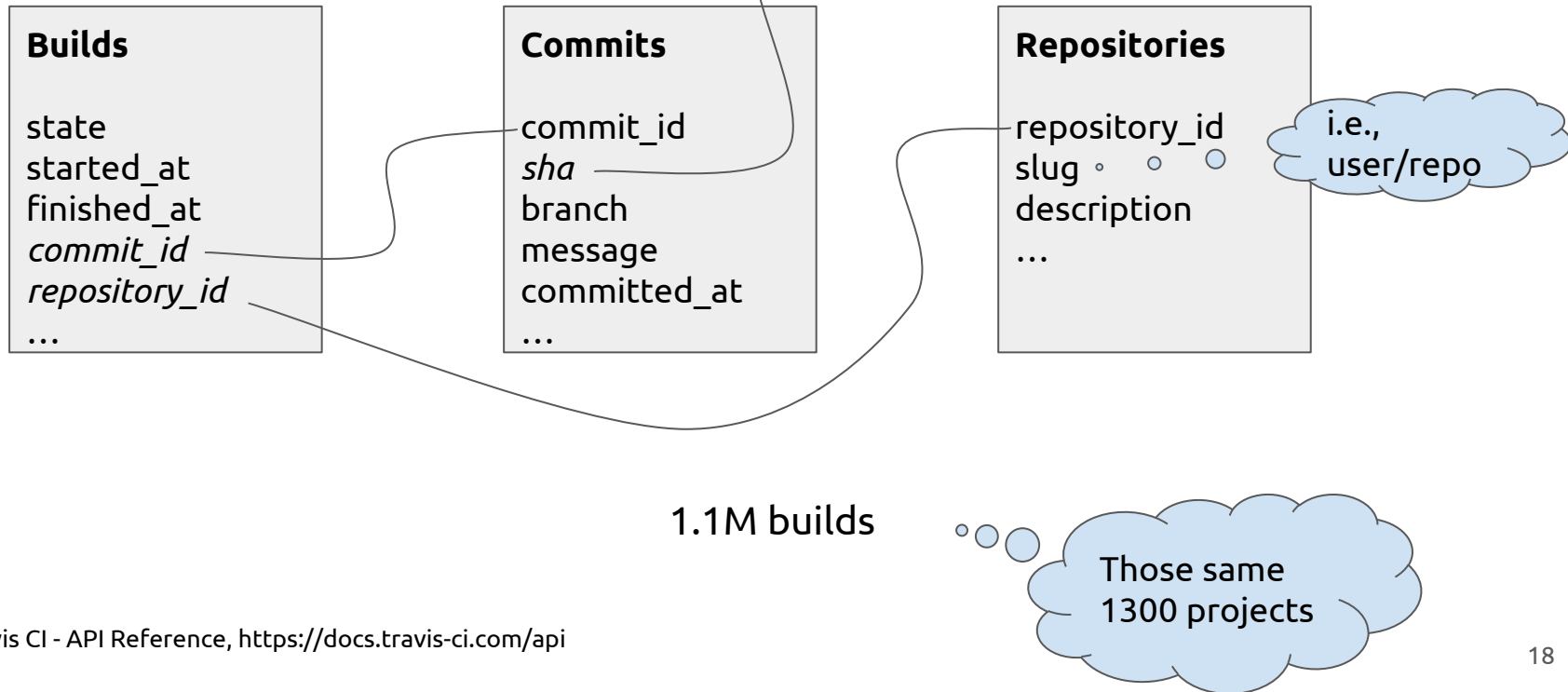
i.e.,
user/repo

Travis CI API

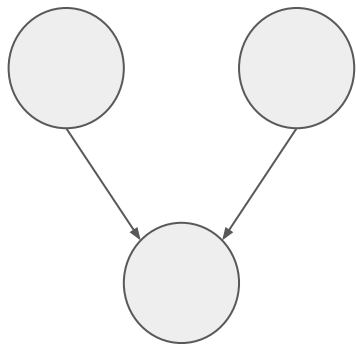
GitHub



GitHub
150GB MySQL dump, 400M commits
Georgios Gionas, The GitHub
Dataset and Tool Suite, MSR 2013.



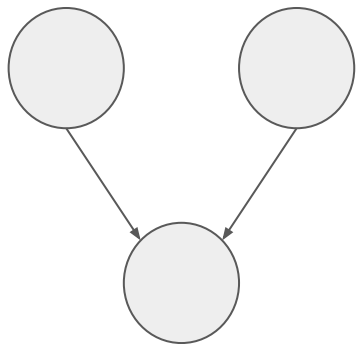
Identify Merge Commits



Number of parents

```
SELECT c.sha,dcs.state,dcs.repository_id
FROM (
  SELECT DISTINCT sha,state,repository_id
  FROM WardTravisCommits AS wtc
  INNER JOIN WardTravisBuilds AS wtb ON wtb.commit_id = wtc.id
) AS dcs
INNER JOIN commits AS c ON c.sha = dcs.sha
INNER JOIN commit_parents AS cp ON c.id = cp.commit_id
GROUP BY c.id,dcs.state
HAVING COUNT(*) > 1
```

Identify Merge Commits

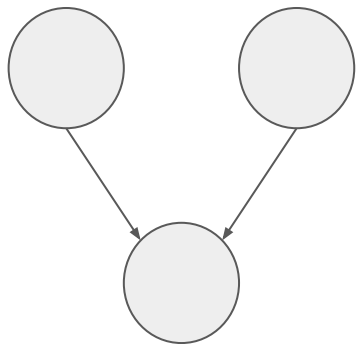


Number of parents

Get commits
and their builds

```
SELECT c.sha,dcs.state,dcs.repository_id
FROM (
  SELECT DISTINCT sha,state,repository_id
  FROM WardTravisCommits AS wtc
  INNER JOIN WardTravisBuilds AS wtb ON wtb.commit_id = wtc.id
) AS dcs
INNER JOIN commits AS c ON c.sha = dcs.sha
INNER JOIN commit_parents AS cp ON c.id = cp.commit_id
GROUP BY c.id,dcs.state
HAVING COUNT(*) > 1
```

Identify Merge Commits



Number of parents

Get commits
and their builds

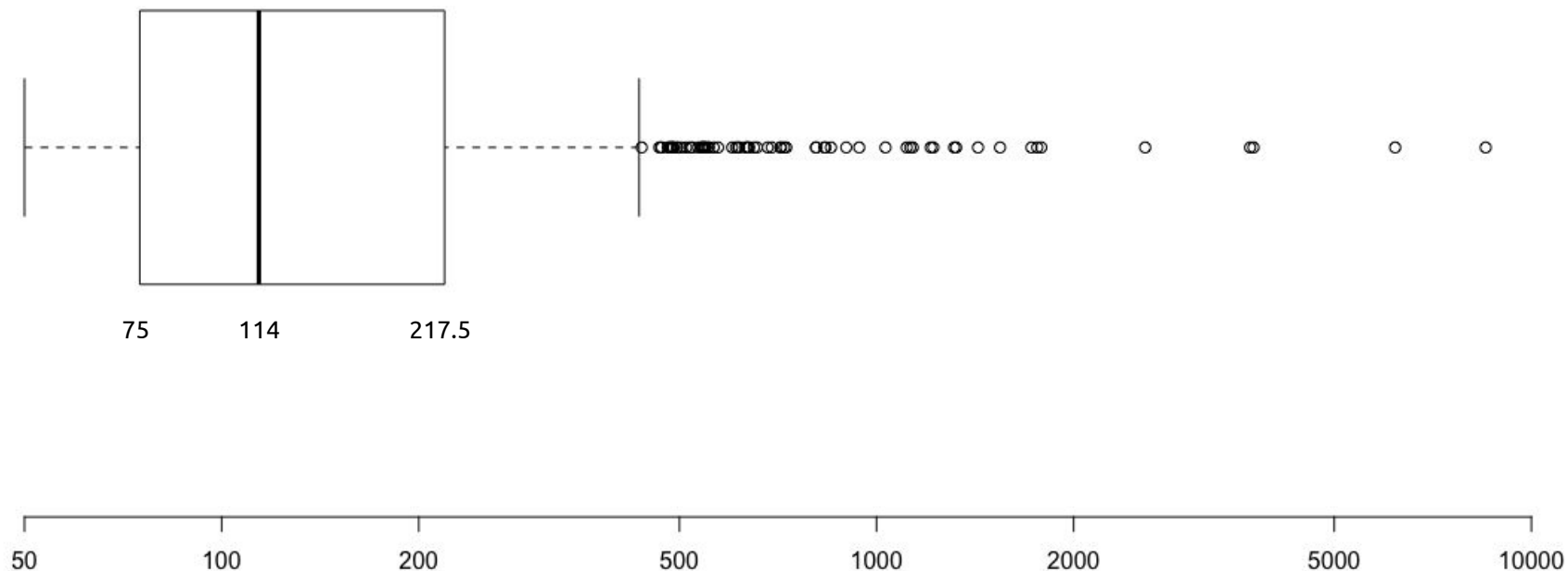
```
SELECT c.sha,dcs.state,dcs.repository_id
FROM (
  SELECT DISTINCT sha,state,repository_id
  FROM WardTravisCommits AS wtc
  INNER JOIN WardTravisBuilds AS wtb ON wtb.commit_id = wtc.id
) AS dcs
```

As long as they
have two or more
parents

```
INNER JOIN commits AS c ON c.sha = dcs.sha
INNER JOIN commit_parents AS cp ON c.id = cp.commit_id
GROUP BY c.id,dcs.state
HAVING COUNT(*) > 1
```

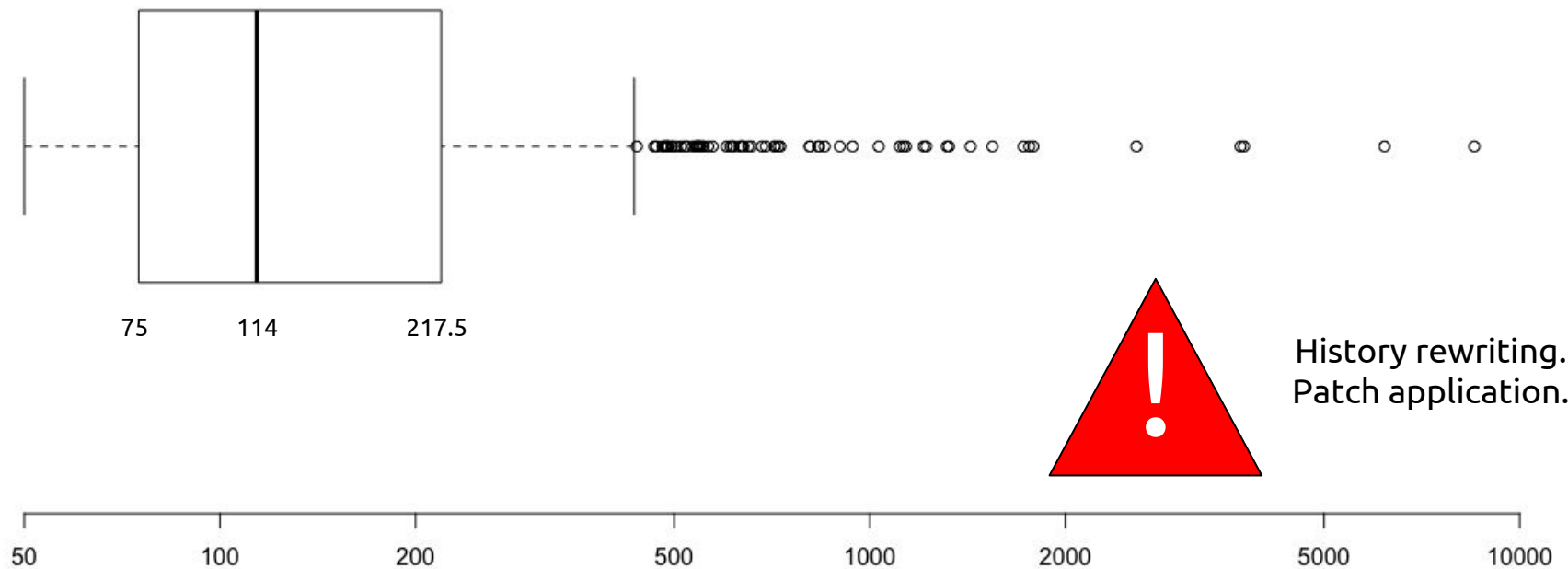
Enough Merge Commits?

584 Projects with 50+ Builds of Merge Commits

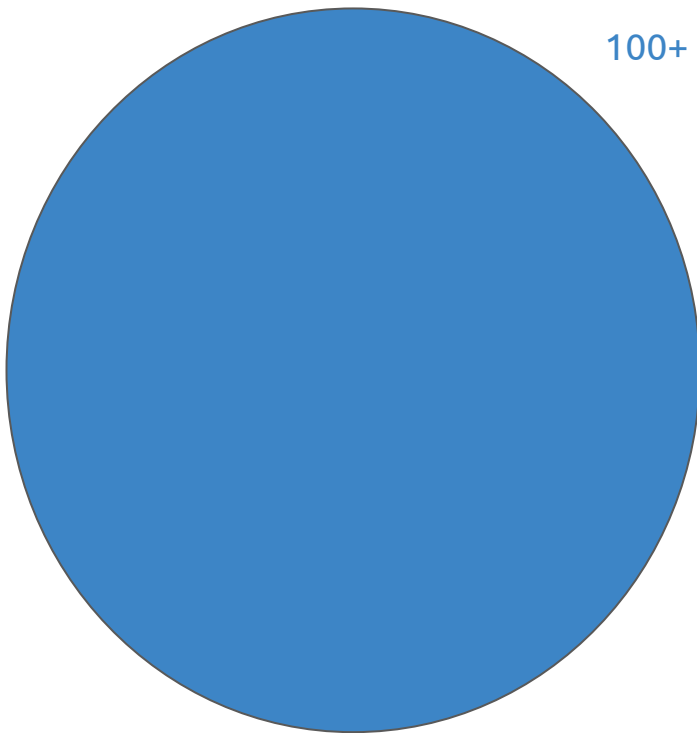


Enough Merge Commits?

584 Projects with 50+ Builds of Merge Commits

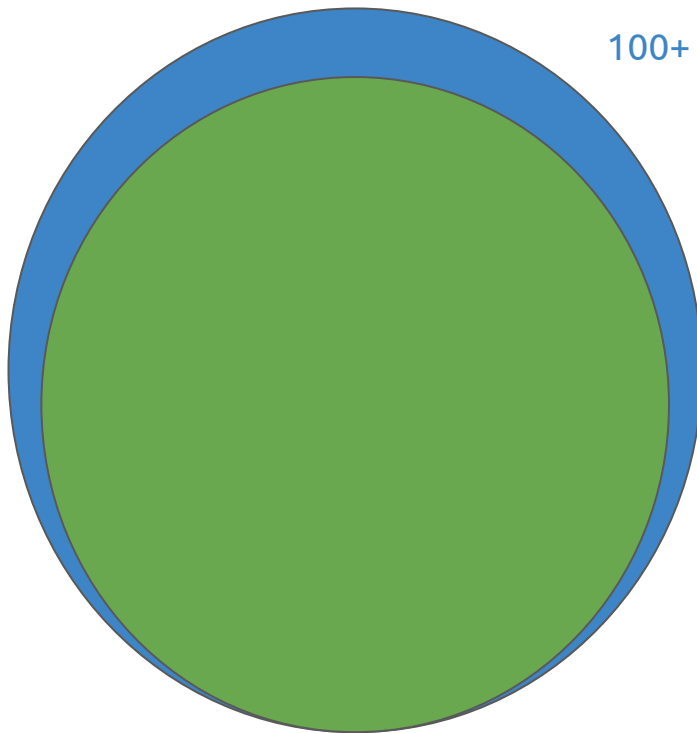


Dataset



100+ builds, 50+ stars, recent commits: 1300 projects

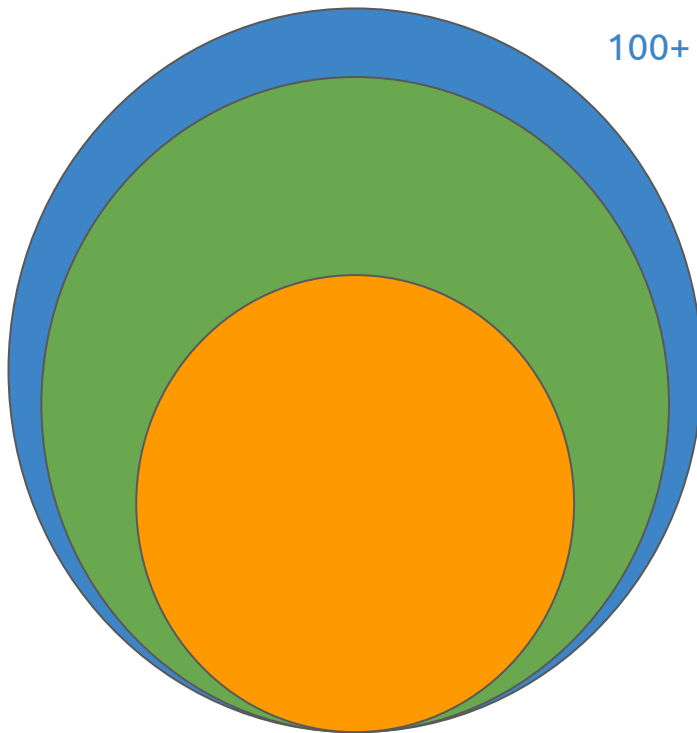
Dataset



100+ builds, 50+ stars, recent commits: 1300 projects

Still on Travis CI: 1248 projects

Dataset







100+ builds, 50+ stars, recent commits: 1300 projects

Still on Travis CI: 1248 projects

50+ builds of merge commits: 584 projects

What Is Failure?

<div>✗ master</div> <div> Brandon Fish</div>	[Truffle] Fix Range#{inspect,to_s} taint specs	<div>✗ #13717 failed</div> <div> 1f6820a</div>	<div>🕒 5 hrs 42 min 2 sec</div> <div>📅 about 18 hours ago</div>
<div>✓ truffle-array-inspect</div> <div> Brandon Fish</div>	[Truffle] Refactoring Array#inspect updates for PR	<div>✓ #13715 passed</div> <div> a7ba831</div>	<div>🕒 5 hrs 40 min 10 sec</div> <div>📅 about 20 hours ago</div>



Travis CI

What Is Failure?

<div>✗ master</div> <div>Brandon Fish</div>	[Truffle] Fix Range#{inspect,to_s} taint specs	<div>✗ #13717 failed</div> <div>1f6820a</div>	<div>🕒 5 hrs 42 min 2 sec</div> <div>📅 about 18 hours ago</div>
<div>✓ truffle-array-inspect</div> <div>Brandon Fish</div>	[Truffle] Refactoring Array#inspect updates for PR	<div>✓ #13715 passed</div> <div>a7ba831</div>	<div>🕒 5 hrs 40 min 10 sec</div> <div>📅 about 20 hours ago</div>

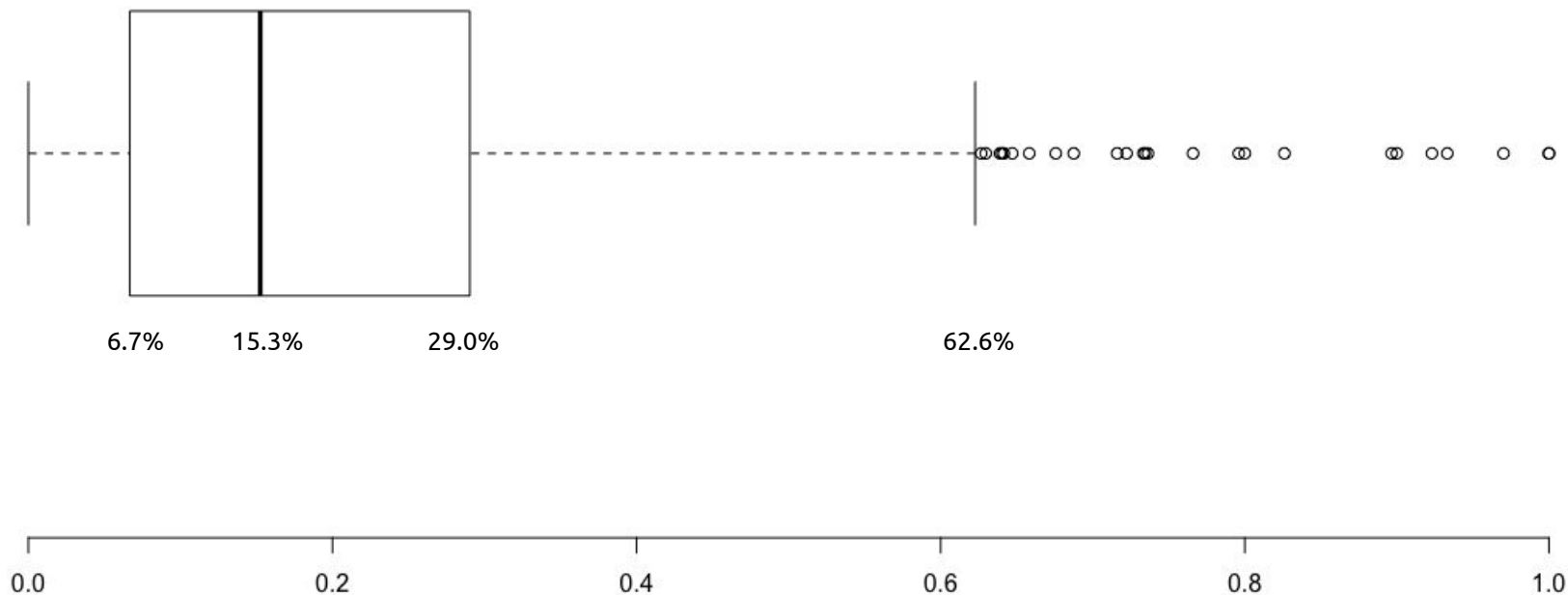


Travis CI

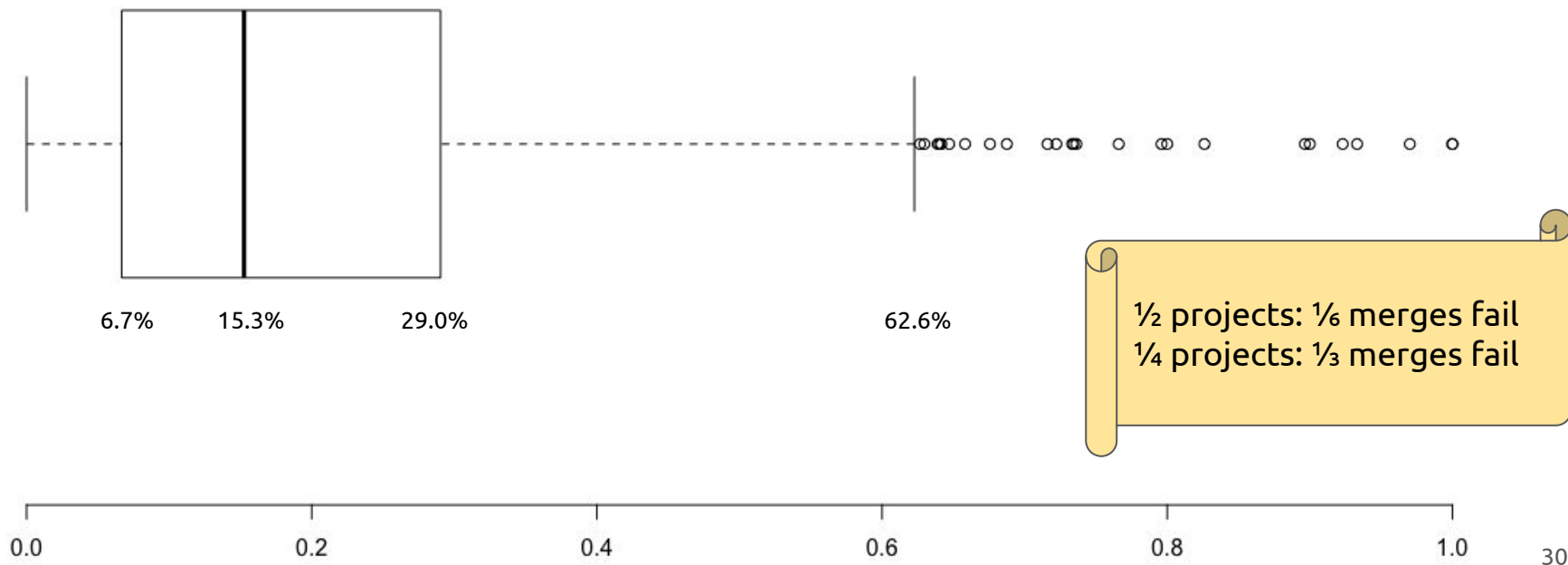
✓ passed
✗ failed
✗ errored

cancelled
started

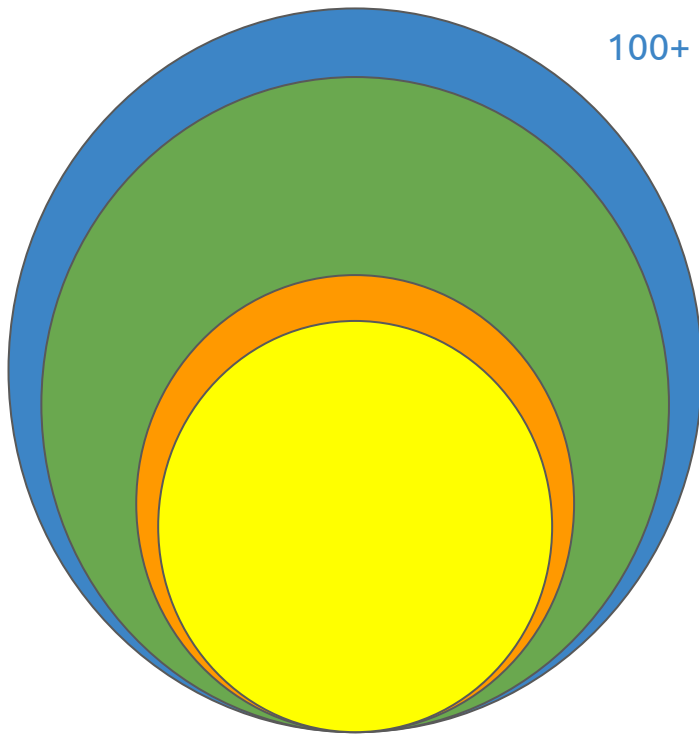
Failed Merge Commit Builds per Project



Failed Merge Commit Builds per Project



Dataset revisited



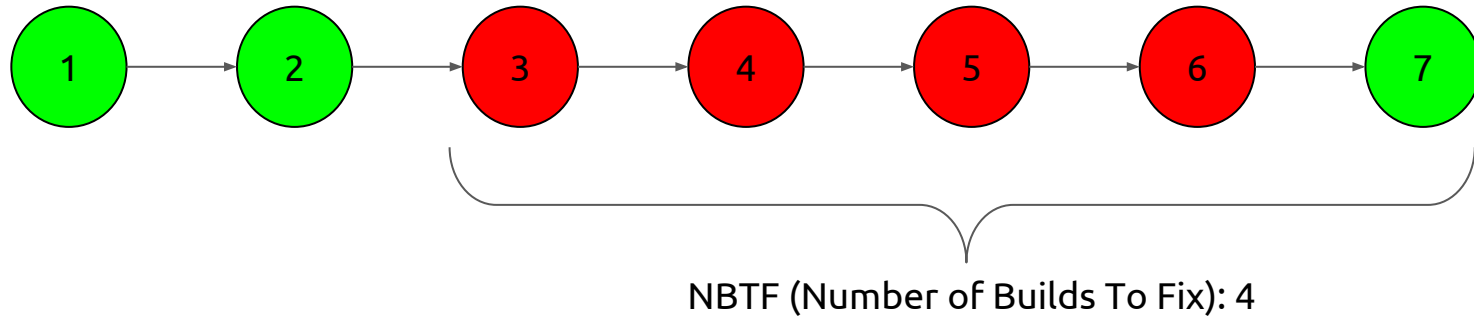
100+ builds, 50+ stars, recent commits: 1300 projects

Still on Travis CI: 1248 projects

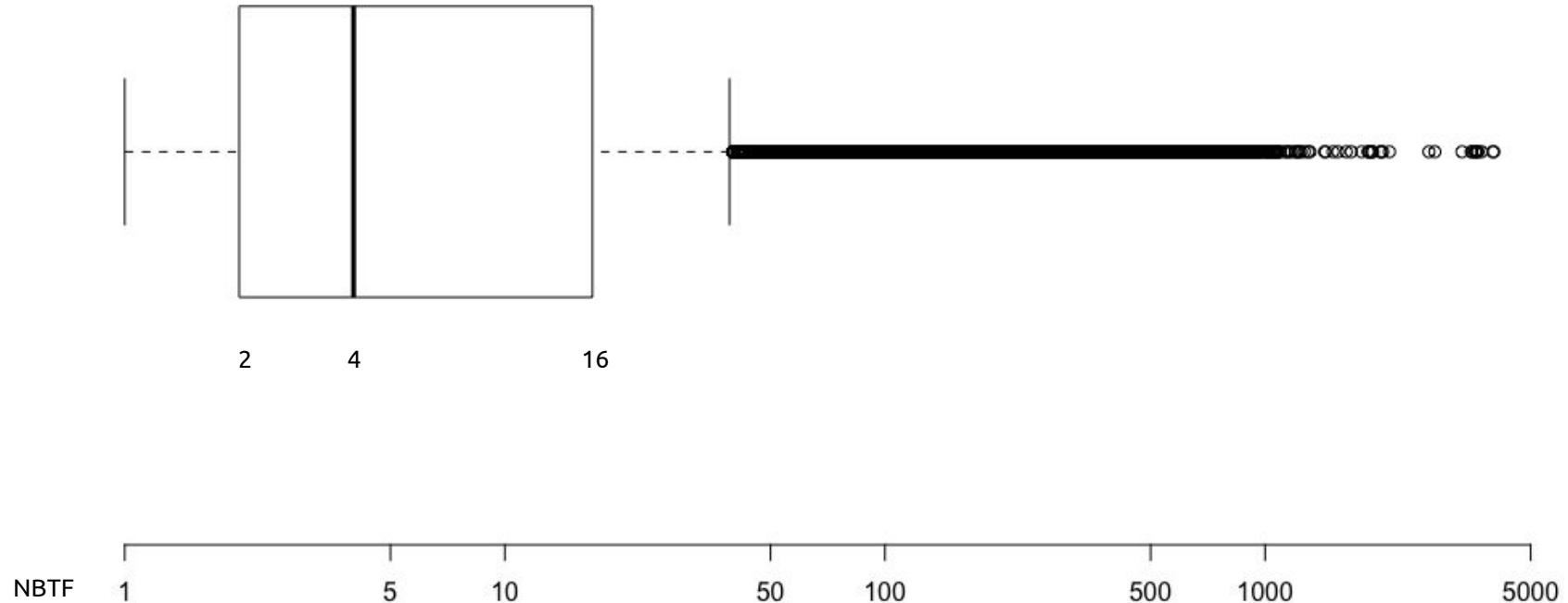
50+ builds of merge commits: 584 projects

Without outliers: 559 projects

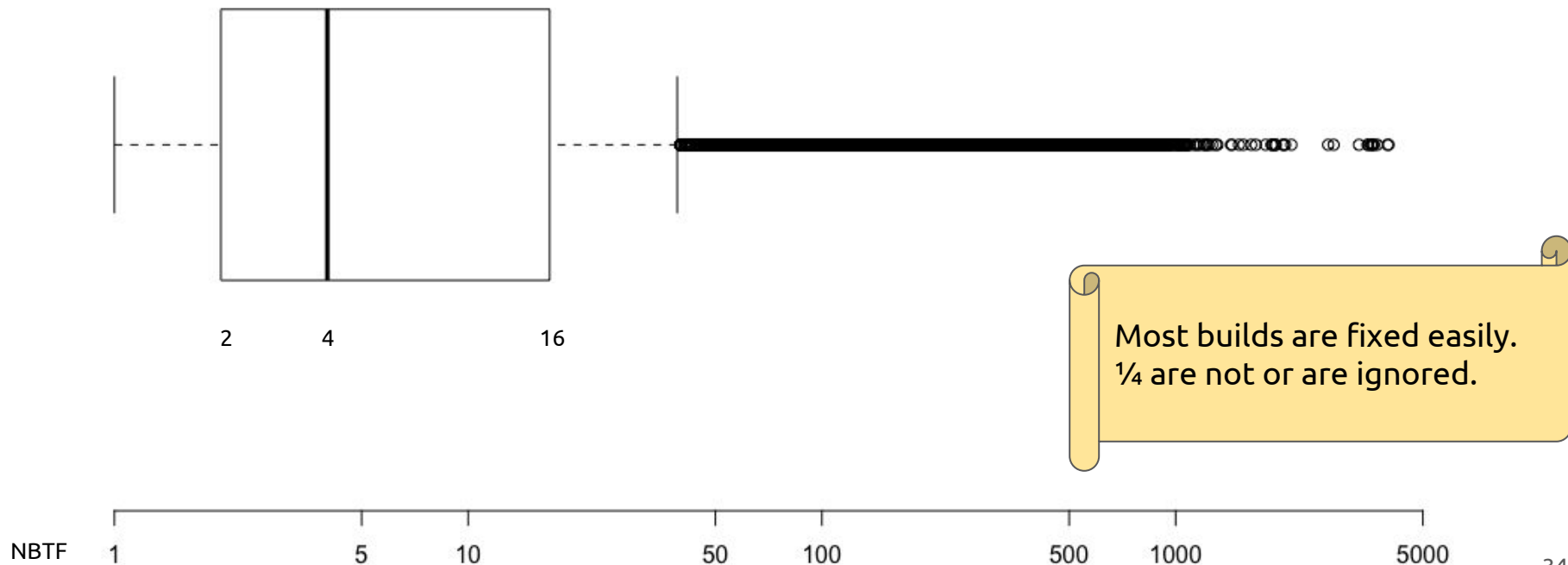
How Much Effort in Fixing the Build?



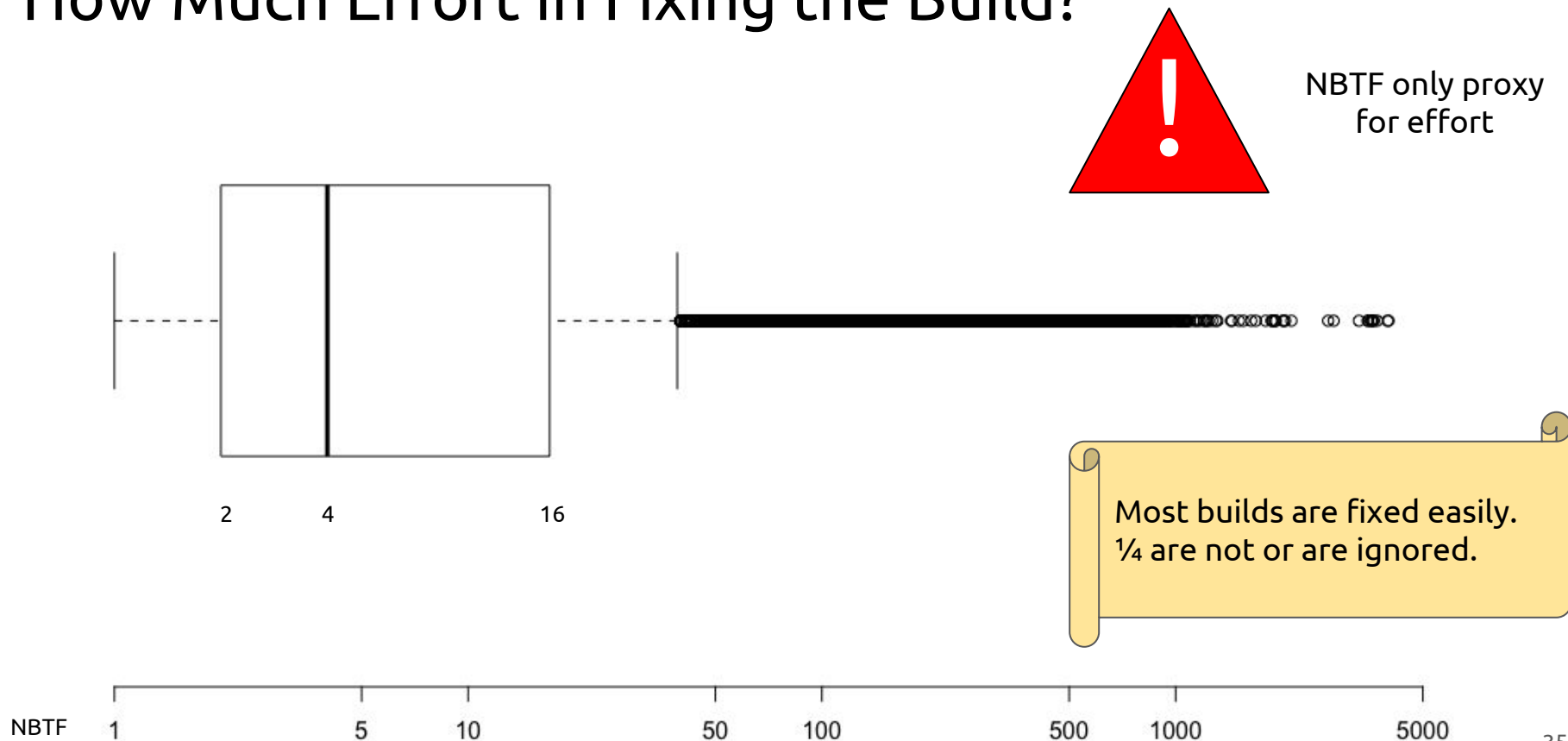
How Much Effort in Fixing the Build?



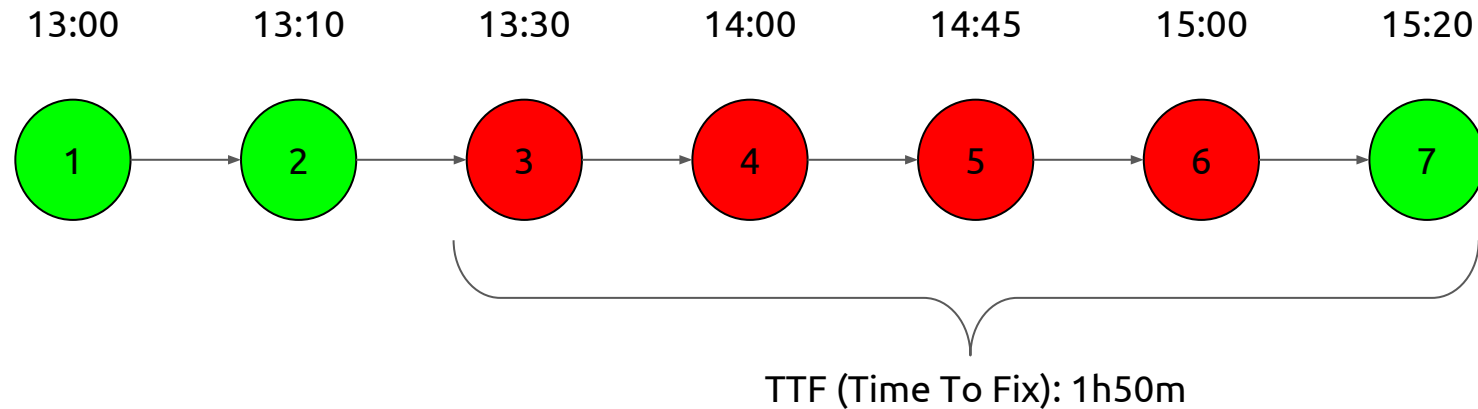
How Much Effort in Fixing the Build?



How Much Effort in Fixing the Build?



How Long to Fix the Build?

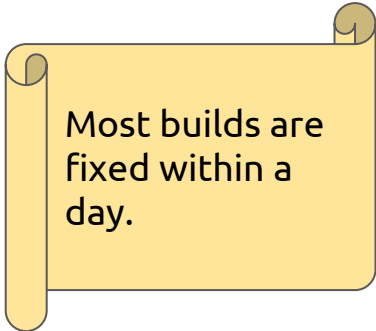


How Long to Fix the Build?

Time range (disjunct)	Amount	Relative	Cumulative
Less than 1 hour	8031	28.2%	28.2%
Less than 12 hours	6062	21.3%	49.4%
Less than 1 day	2392	8.4%	57.8%
Less than 7 days	5556	19.5%	77.3%
Less than 1 month (30 days)	3321	11.6%	88.9%
Less than 1 year (365 days)	2967	10.4%	99.3%
More than 1 year	198	0.7%	100%

How Long to Fix the Build?

Time range (disjunct)	Amount	Relative	Cumulative
Less than 1 hour	8031	28.2%	28.2%
Less than 12 hours	6062	21.3%	49.4%
Less than 1 day	2392	8.4%	57.8%
Less than 7 days	5556	19.5%	77.3%
Less than 1 month (30 days)	3321	11.6%	88.9%
Less than 1 year (365 days)	2967	10.4%	99.3%
More than 1 year	198	0.7%	100%



Most builds are fixed within a day.

How Long to Fix the Build?

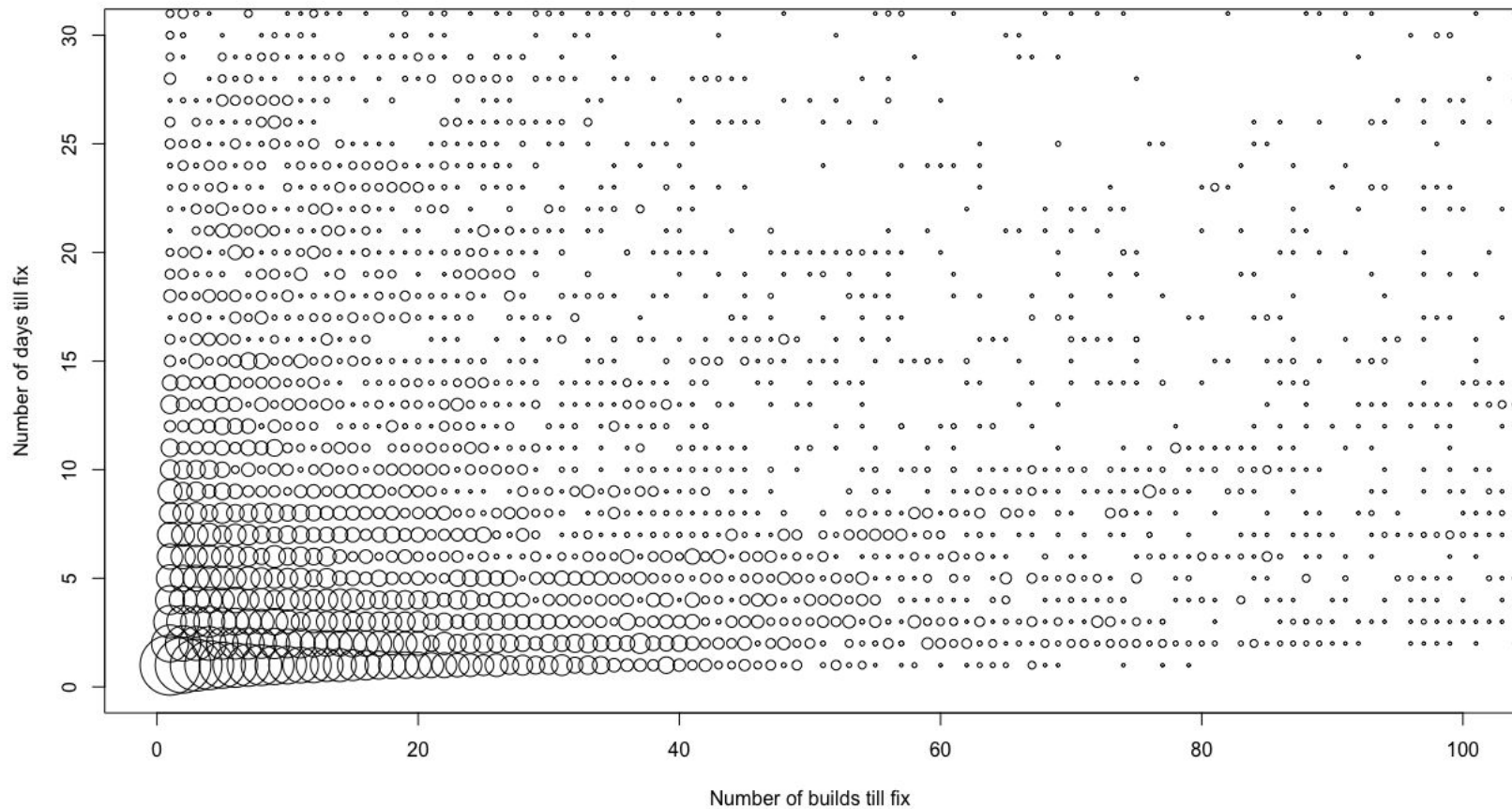
Time range (disjunct)	Amount	Relative	Cumulative
Less than 1 hour	8031	28.2%	28.2%
Less than 12 hours	6062	21.3%	49.4%
Less than 1 day	2392	8.4%	57.8%
Less than 7 days	5556	19.5%	77.3%
Less than 1 month (30 days)	3321	11.6%	88.9%
Less than 1 year (365 days)	2967	10.4%	99.3%
More than 1 year	198	0.7%	100%



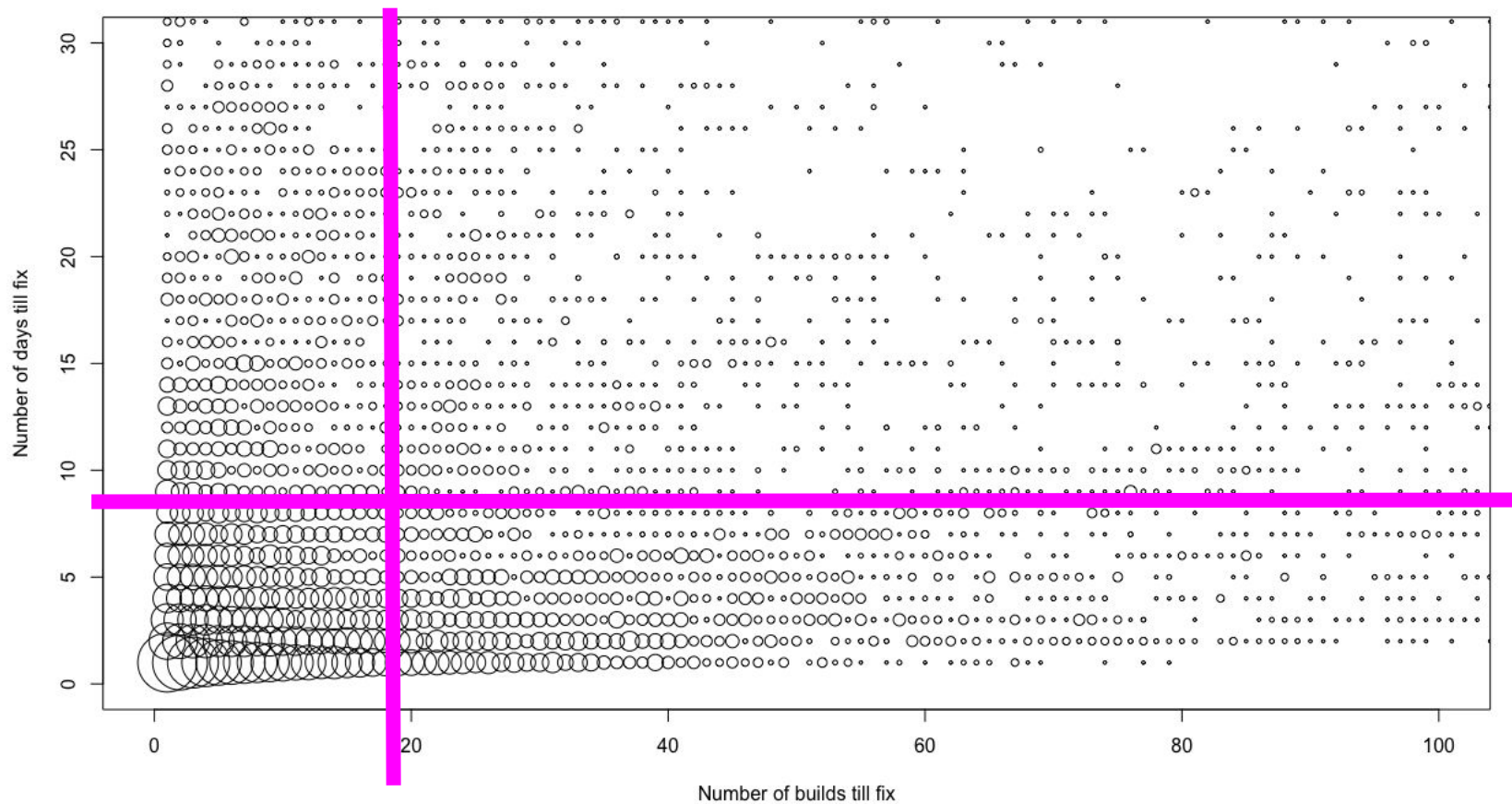
Time taken
≠
time worked
on

Most builds are
fixed within a
day.

Number of Builds To Fix vs Time To Fix



Number of Builds To Fix vs Time To Fix



Answers to Research Questions

RQ1

Conflicts after code integration happen often.

For half of the projects, the build did not succeed for 1/6 merge commits.

RQ2

Multiple builds are needed to fix the majority of botched code integrations.

RQ3

The majority of botched code integrations are fixed within a day.

Future

This study

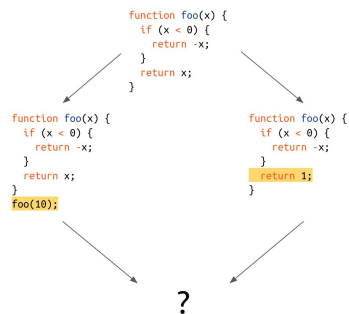
- Filter out projects that do not *really* follow continuous delivery
- Refine effort estimation using commit information
- Use TravisTorrent information to refine failure categorization

Long term

- Help developers with merging
- Behavioural differencing with dynamic symbolic execution

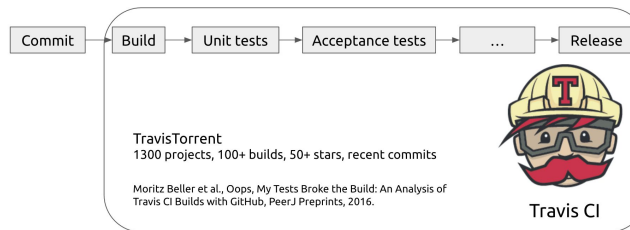
Summary

Semantic Conflicts



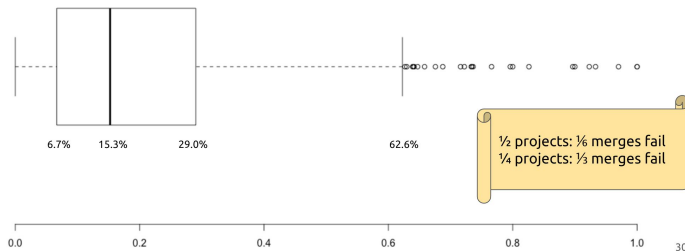
7

Continuous Delivery



13

Failed Merge Commit Builds per Project



Future

This study

- Filter out projects that do not *really* follow continuous delivery
- Refine effort estimation using commit information
- Use TravisTorrent information to refine failure categorization

Long term

- Help developers with merging
- Behavioural differencing with dynamic symbolic execution

43