Broadcast vs. Unicast Review Technology: Does it Matter?

Foutse Khomh
foutse.khomh@polymtl.ca
@SWATLab

56th CREST Open Workshop

SWAT

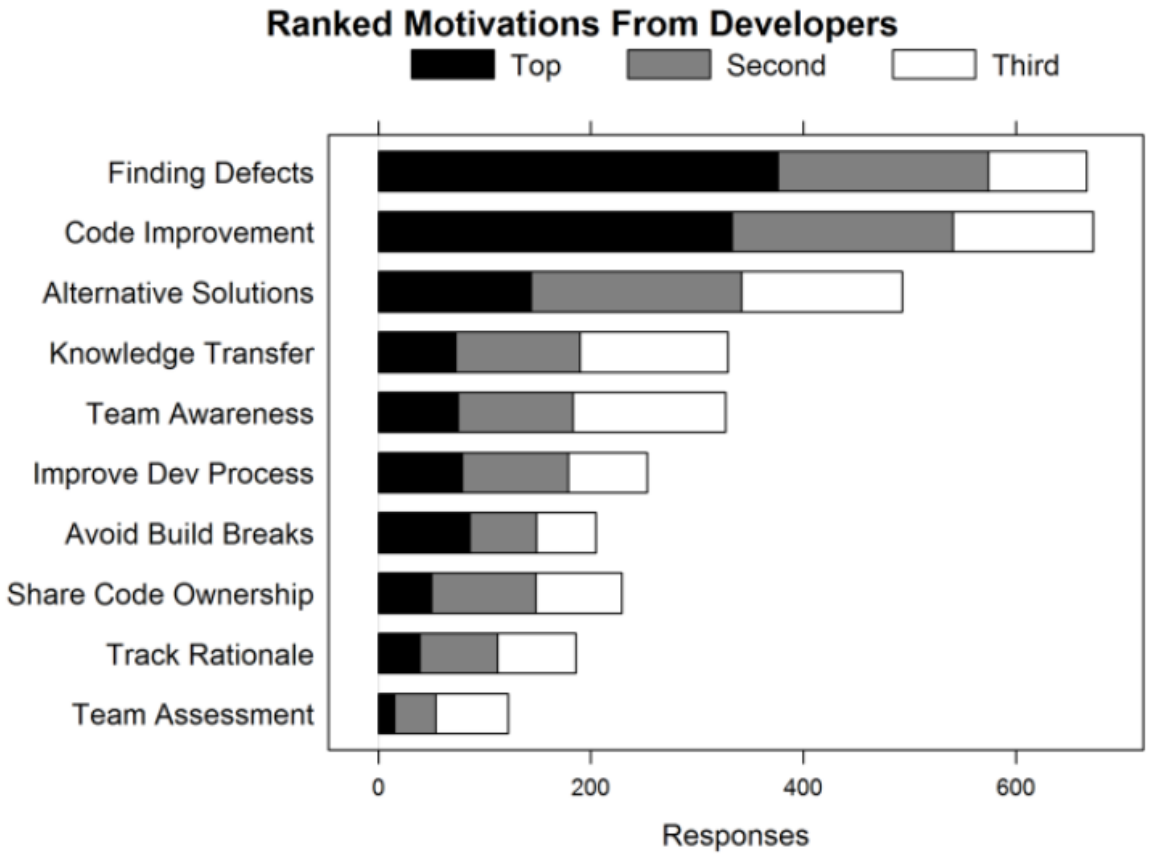
POLYTECHNIQUE MONTRÉAL
WORLD-CLASS ENGINEERING

Code review:

What is in that name?

Code review:

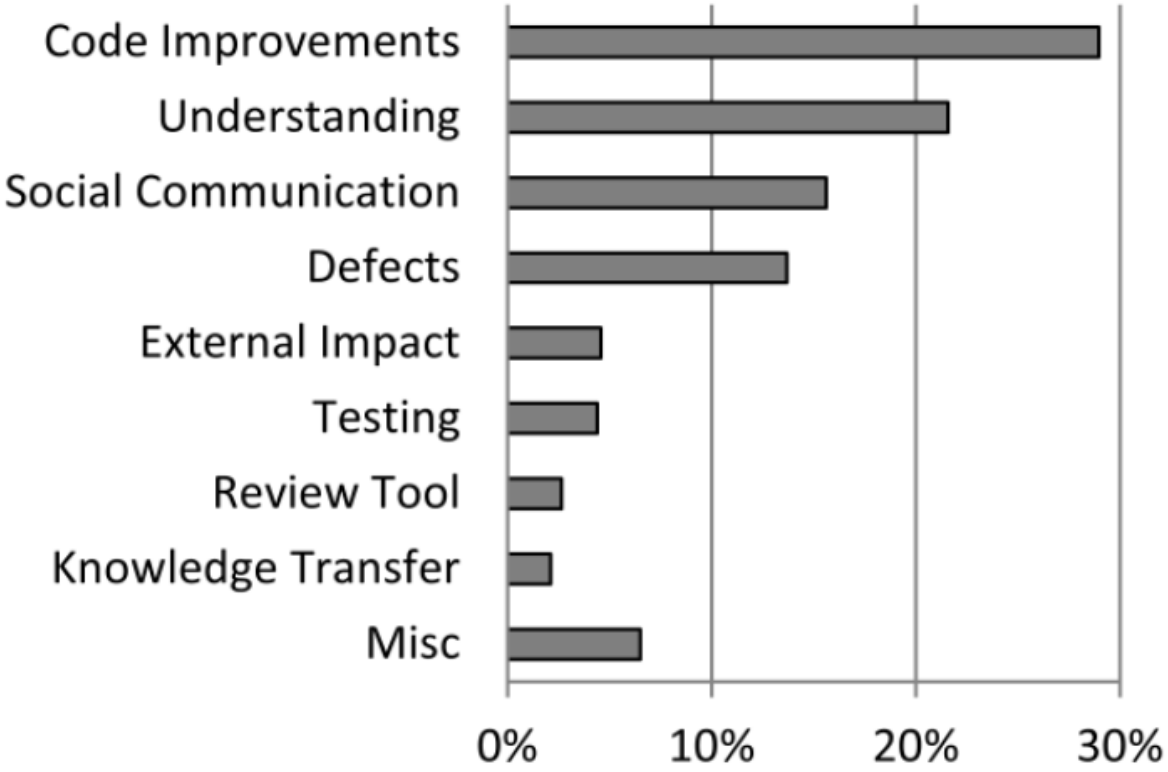
Why?



"Expectations, Outcomes, and Challenges of Modern Code Review"
Alberto Bacchelli and Christian Bird - ICSE 2013

Code review:

Common outcomes …



"Expectations, Outcomes, and Challenges of Modern Code Review"
Alberto Bacchelli and Christian Bird - ICSE 2013

Code review:

Different types …

Over-the-Shoulder

The Email Thread

Tool-Assisted

Pair Programming

Code review:

Different types …

Over-the-Shoulder

The Email Thread

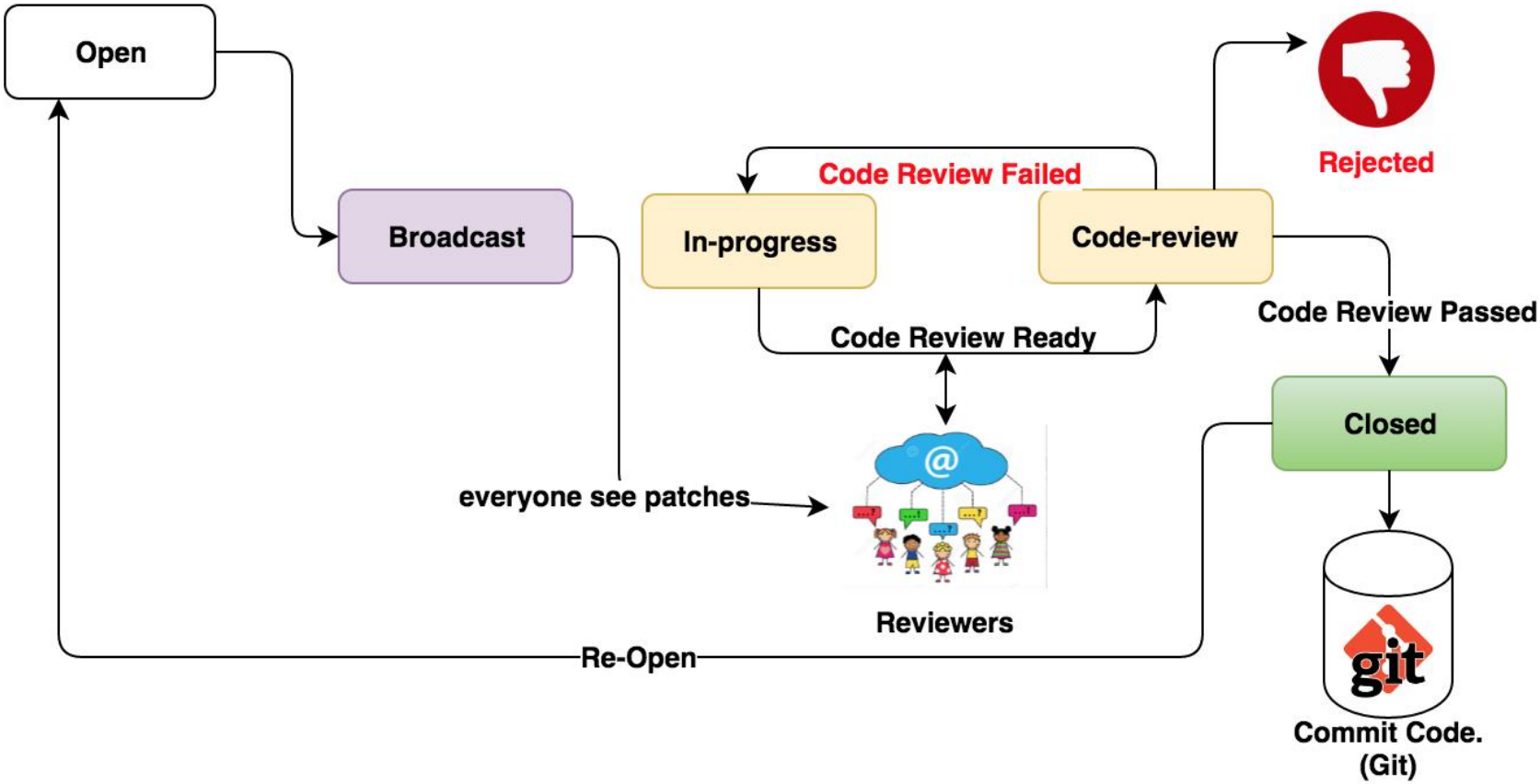
broadcast

Tool-Assisted

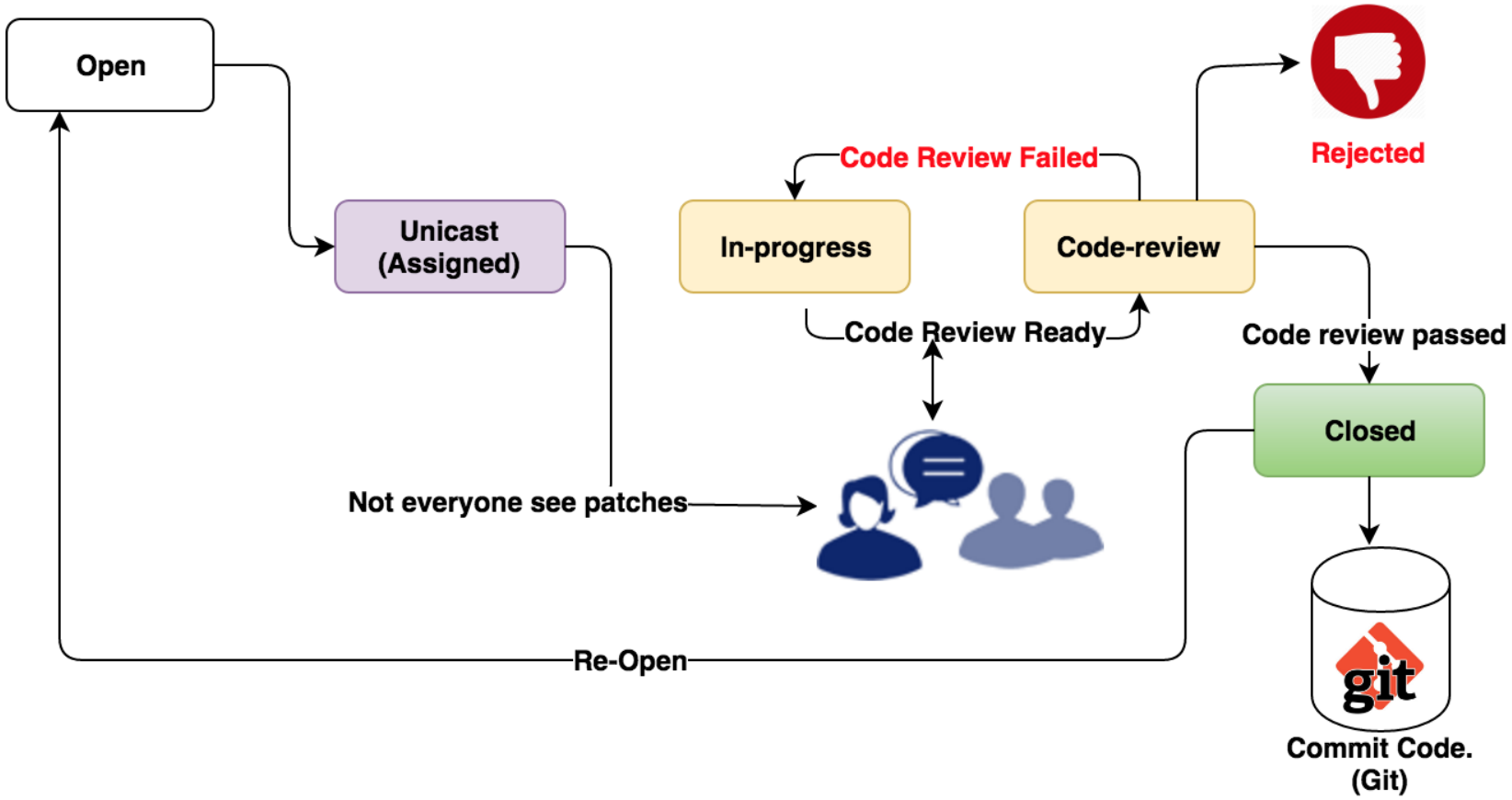
Unicast

Pair Programming

Work-flow and Status in Broadcast Environment



Work-flow and Status in Unicast Environment



Does the Medium Technology Used for Code Reviews Affects Reviews Activities

Effort

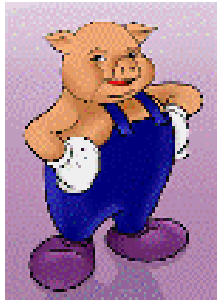
Effectiveness

Efficiency

Apache Pig

Case Studies

SUBVERSION®

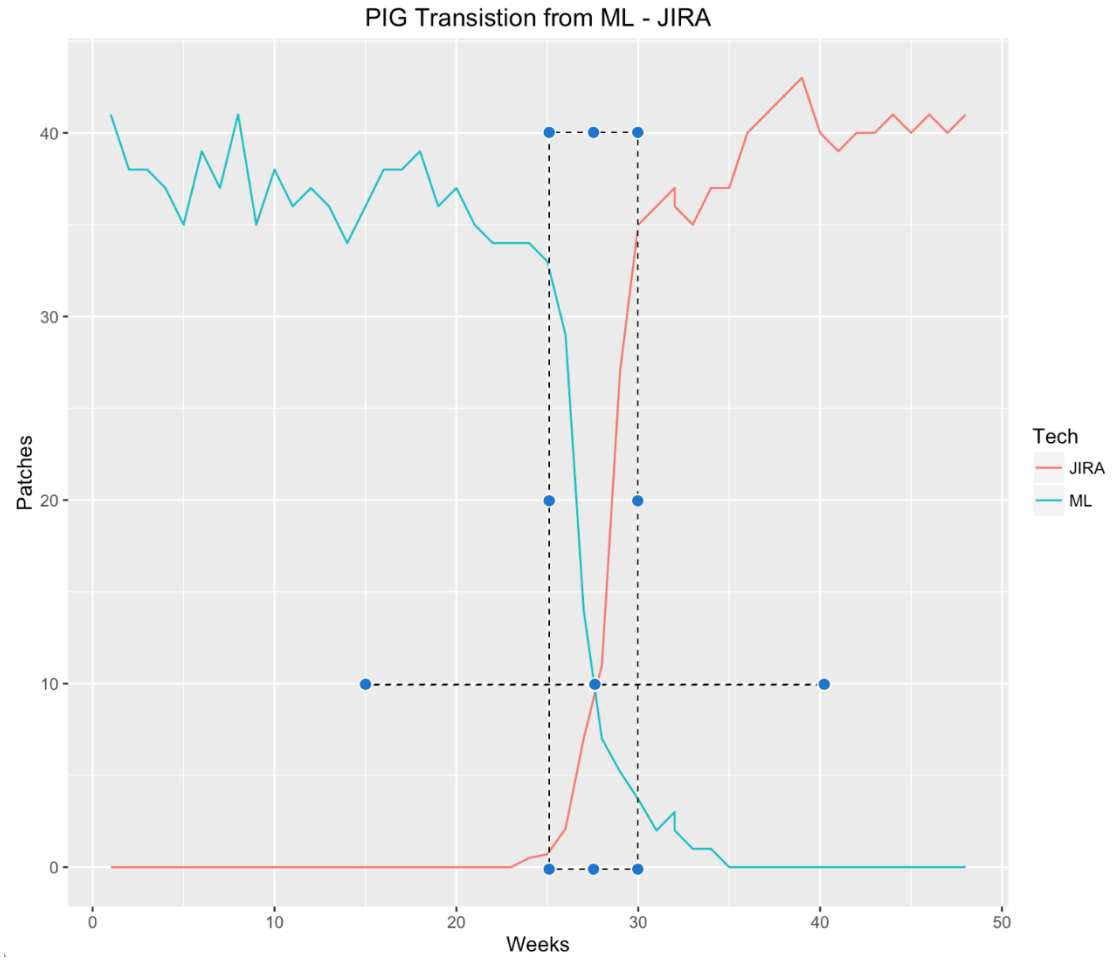


Apache Pig

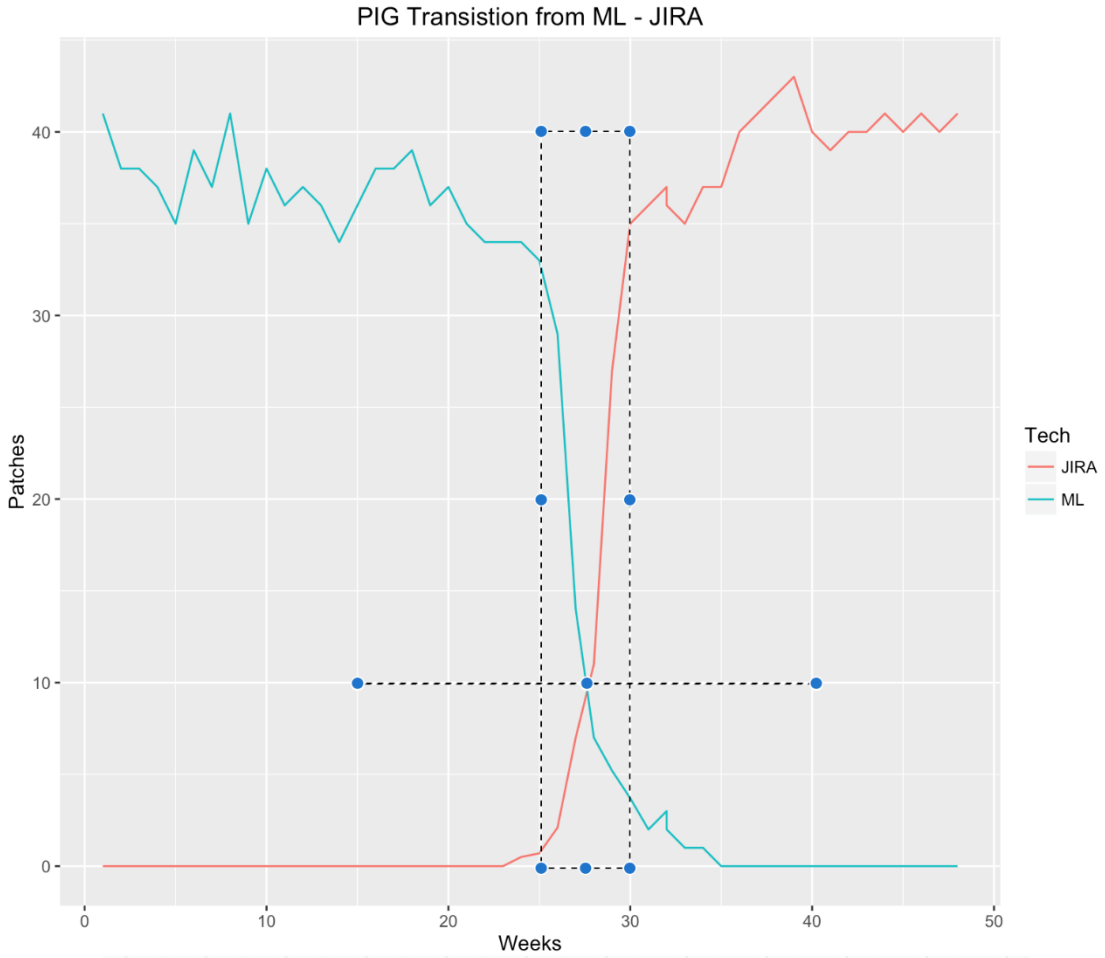
SUBVERSION®

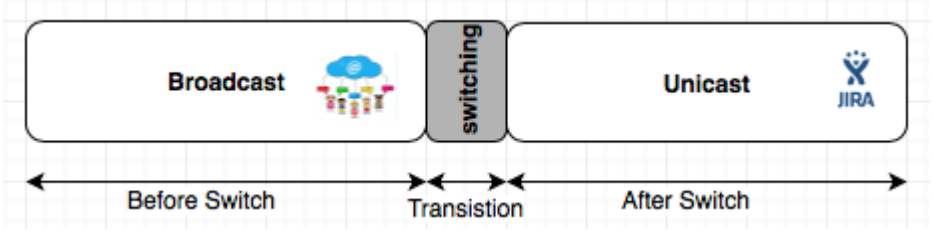
Five Apache projects that transitioned from Broadcast to Unicast technology

Identification of the transition period



Identification of the transition period

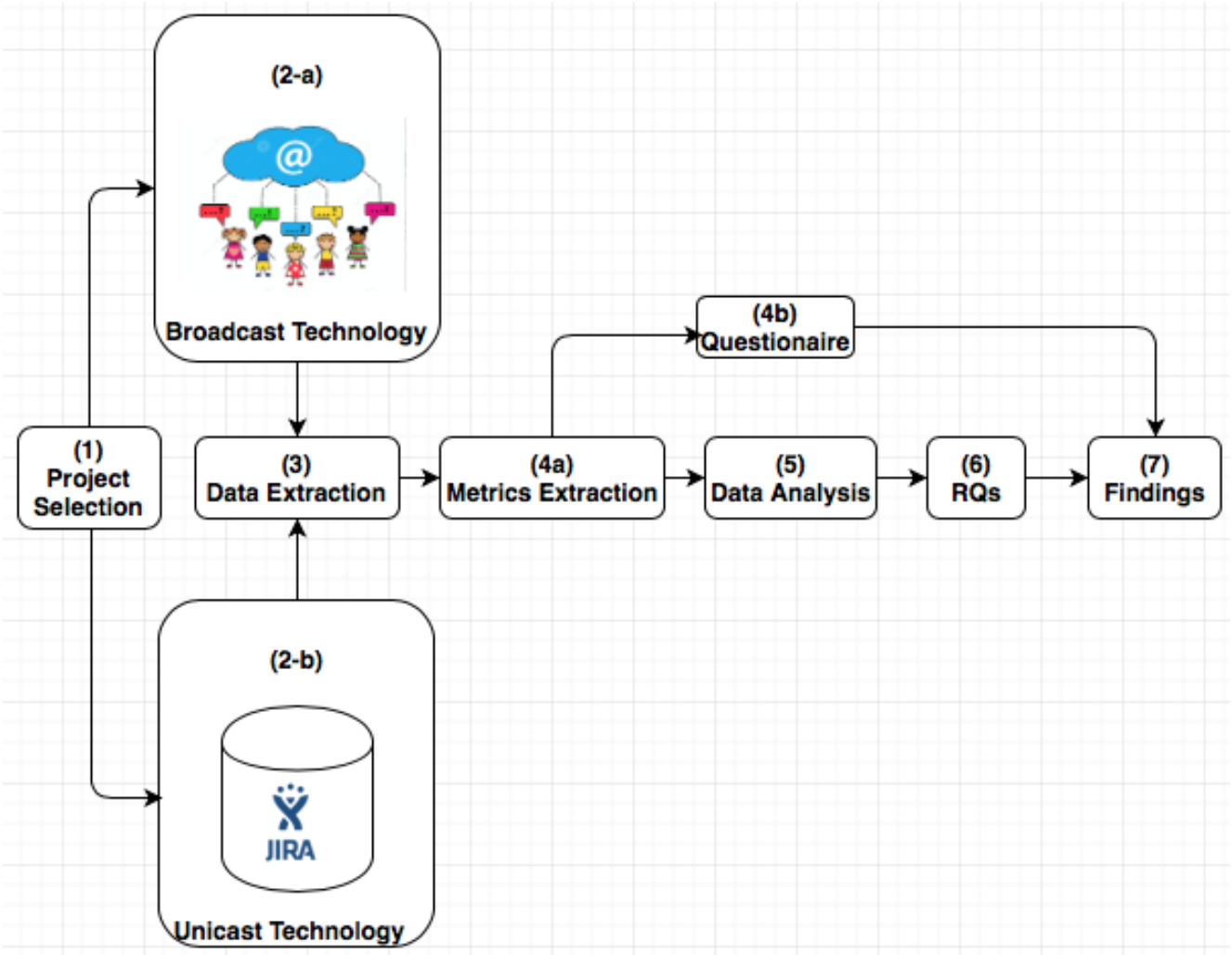




Research Questions

- **RQ1: Is review effort related to the review medium used?**
- **RQ2: Is the effectiveness of a patch reviewing process related to the medium used?**
- **RQ3: Is the efficiency of a patch review process related to the medium used?**

Data Extraction



RQ1: Is review effort related to the review medium used?

- ❑ Number of developers involved in a review (NV)
- ❑ Number of rounds necessary to review a patch (NR)
- ❑ Number of review requests for a patch (RQu)

Approach

- **H0 :** ***There is no significant difference between the value of metric m for patches reviewed on broadcasts and those reviewed on unicast***

❑ **Mann-Whitney U test**

❑ **Cliff's Delta effect Size**

Findings

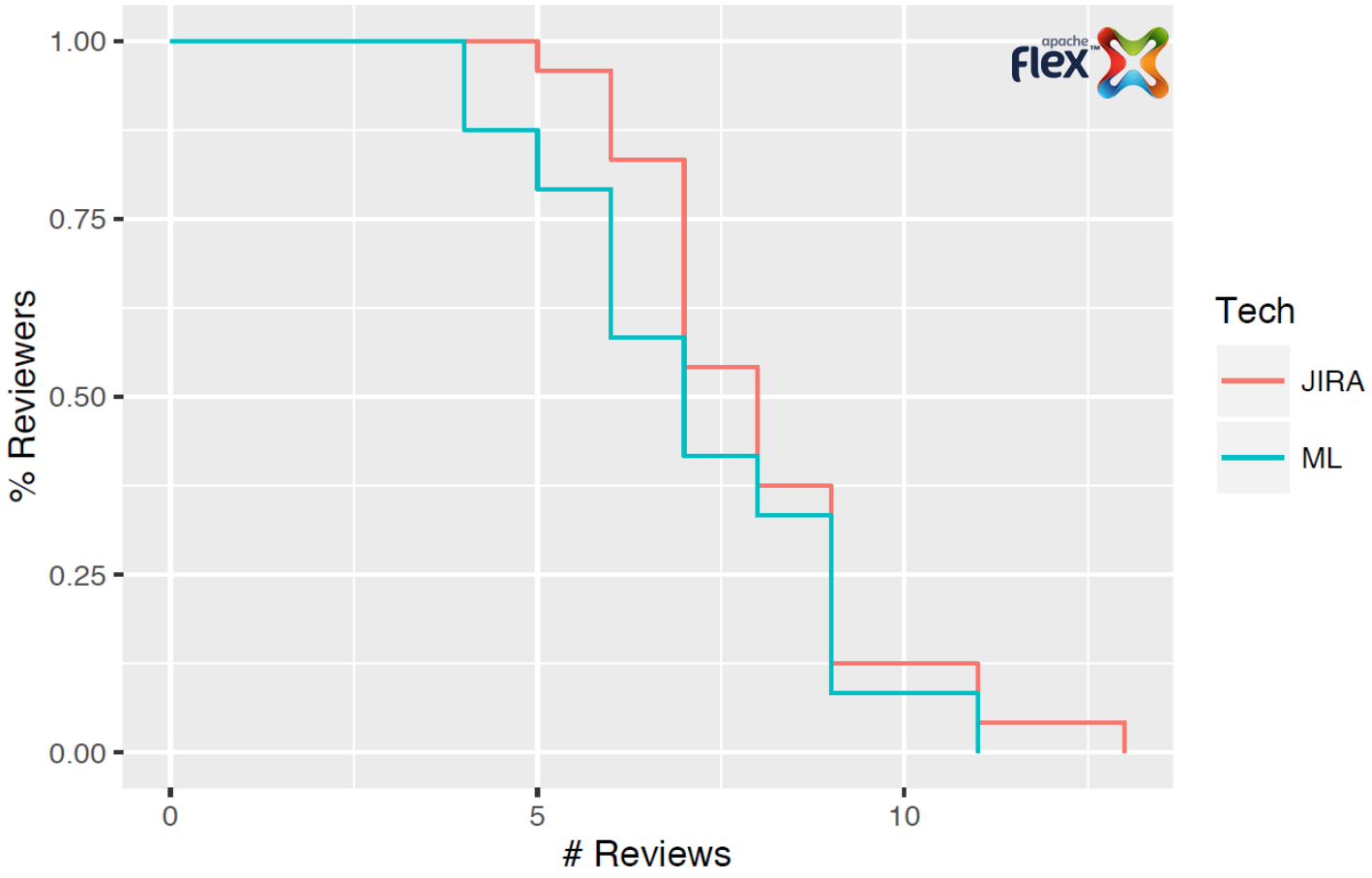
	Number of developers		Number of rounds		Number of requests	
	P-values	Cliff's d	P-values	Cliff's d	P-values	Cliff's d
Flex	0.843	Negligible	**0.016**	**Large**	**0.012**	**Large**
HBase	0.891	Negligible	**0.021**	**Large**	**0.017**	**Large**
Hive	0.902	Negligible	**0.011**	**Large**	**0.031**	**Large**
Pig	0.931	Negligible	**0.017**	**Large**	**0.021**	**Large**
Subversion	0.915	Negligible	**0.013**	**Large**	**0.019**	**Large**

Findings

	Number of developers		Number of rounds		Number of requests	
	P-values	Cliff's d	P-values	Cliff's d	P-values	Cliff's d
Flex	0.843	Negligible	**0.016**	**Large**	**0.012**	**Large**
HBase	0.891	Negligible	**0.021**	**Large**	**0.017**	**Large**
Hive	0.902	Negligible	**0.011**	**Large**	**0.031**	**Large**
Pig	0.931	Negligible	**0.017**	**Large**	**0.021**	**Large**
Subversion	0.915	Negligible	**0.013**	**Large**	**0.019**	**Large**

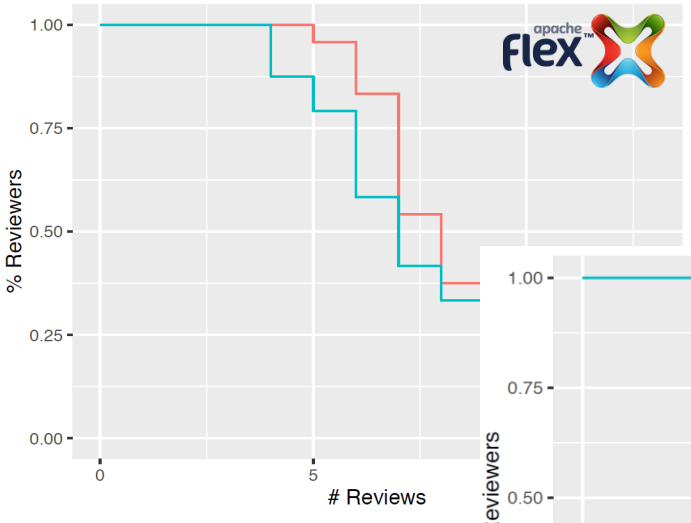
More iterations and more requests on Unicast

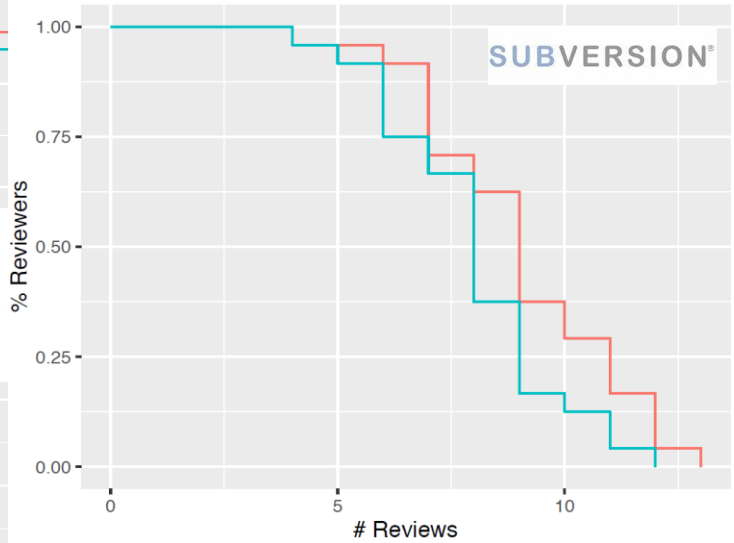
Findings

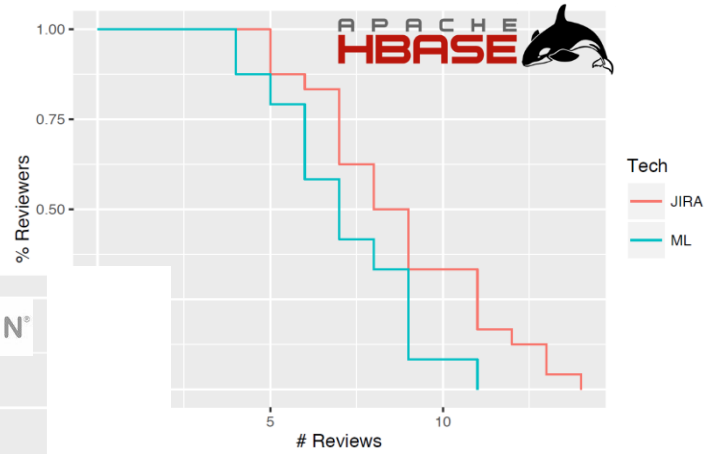


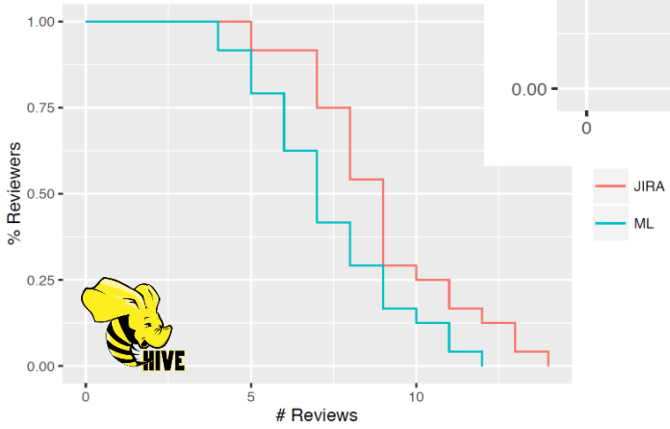

WORK

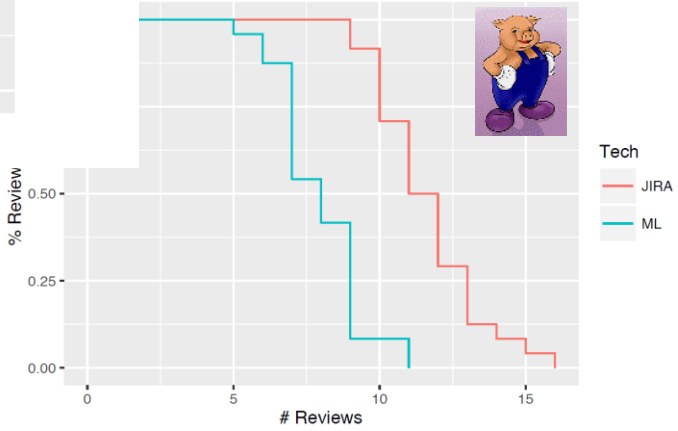
Findings


apache Flex
1.00
0.75
0.50
0.25
0.00
% Reviewers
0
5
Reviews
Tech
JIRA


SUBVERSION
1.00
0.75
0.50
0.25
0.00
% Reviewers
0
5
10
Reviews


APACHE HBASE
1.00
0.75
0.50
% Reviewers
5
10
Reviews
Tech
JIRA
ML


1.00
0.75
0.50
0.25
0.00
% Reviewers
0
5
10
Reviews
JIRA
ML
HIVE


0.50
0.25
0.00
% Review
0
5
10
15
Reviews
Tech
JIRA
ML

RQ1: Is review effort related to the review medium used?

YES

- **Patches reviewed on unicast undergo more iterations, and**
- **Unicast's reviewers are more active during code review.**

RQ2: Is the effectiveness of a patch reviewing process related to the medium used?

- ❑ Post review bugs
- ❑ Median review rate (MRR)

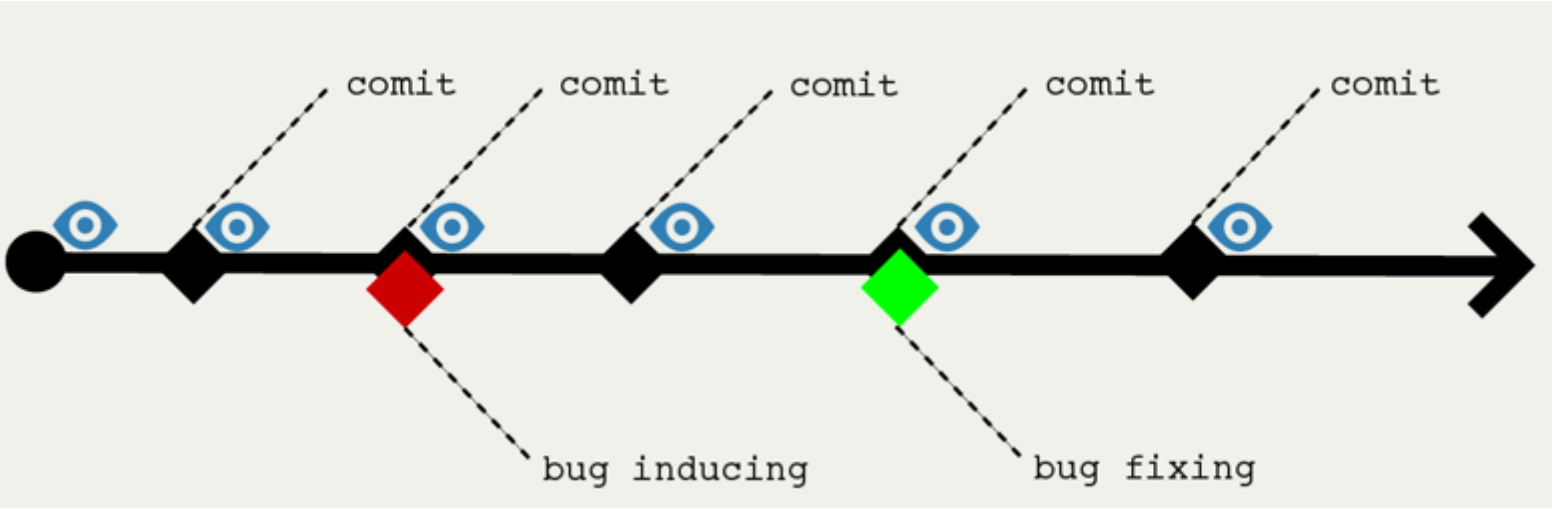
Approach

- **H0 :** ***There is no significant difference between the value of metric m for patches reviewed on broadcasts and those reviewed on unicast***

  - ❑ **Mann-Whitney U test**
  - ❑ **Cliff's Delta effect Size**
  - ❑ **The SZZ Algorithm to link bugs with reviewed patches**

Approach

- **SZZ Algorithm**



Findings

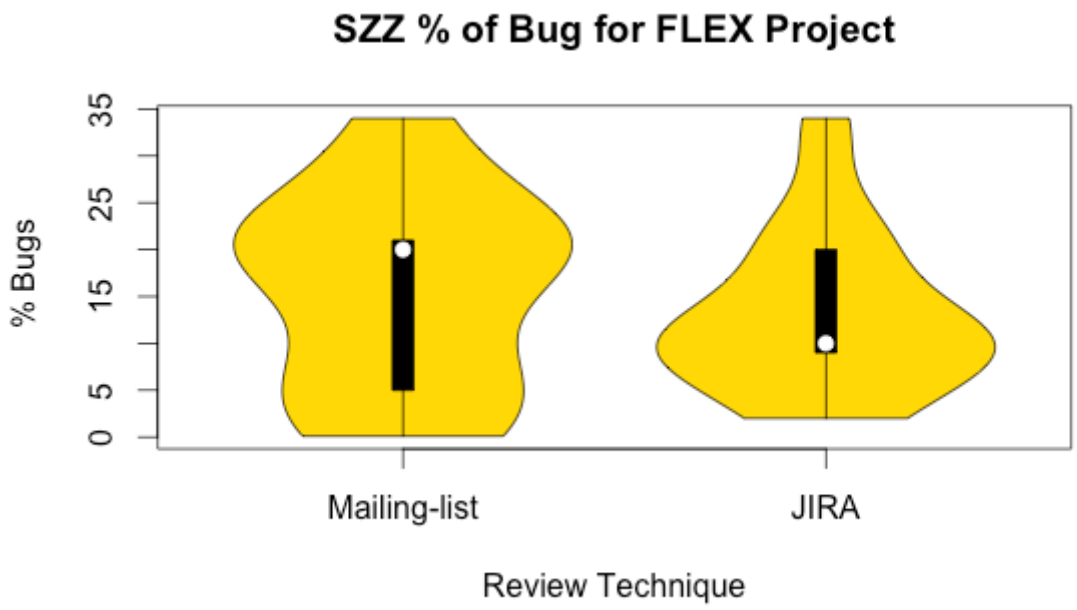
	Post review bugs		Median review rate	
	P-values	Cliff's d	P-values	Cliff's d
Apache Flex	0.015	Medium	0.021	Large
Apache HBase	0.039	Medium	0.017	Large
Apache Hive	0.031	Medium	0.014	Large
Apache Pig	0.017	Medium	0.011	Large
Subversion	0.008	Medium	0.031	Large

Findings

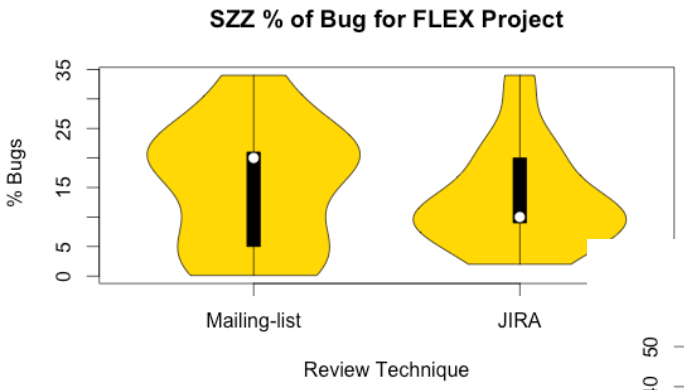
	Post review bugs		Median review rate	
	P-values	Cliff's d	P-values	Cliff's d
Flex	0.015	Medium	0.021	Large
HBase	0.039	Medium	0.017	Large
Hive	0.031	Medium	0.014	Large
Pig	0.017	Medium	0.011	Large
Subversion	0.008	Medium	0.031	Large

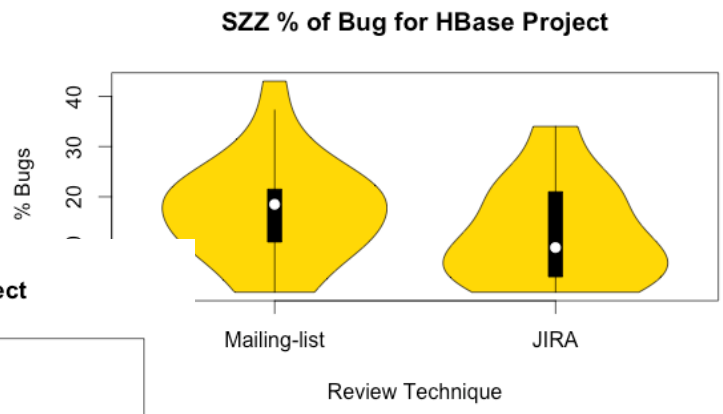
Fewer post review bugs on Unicast and more review activities

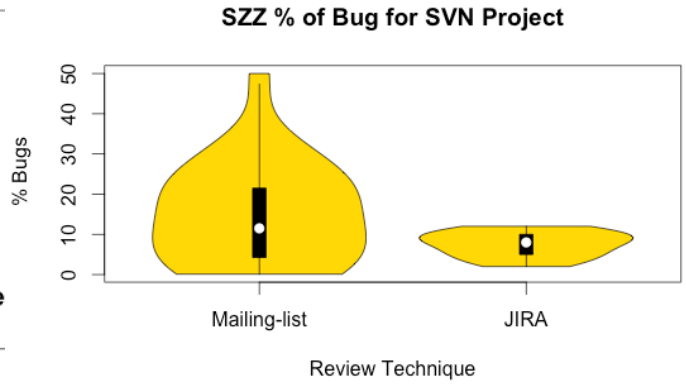
Findings

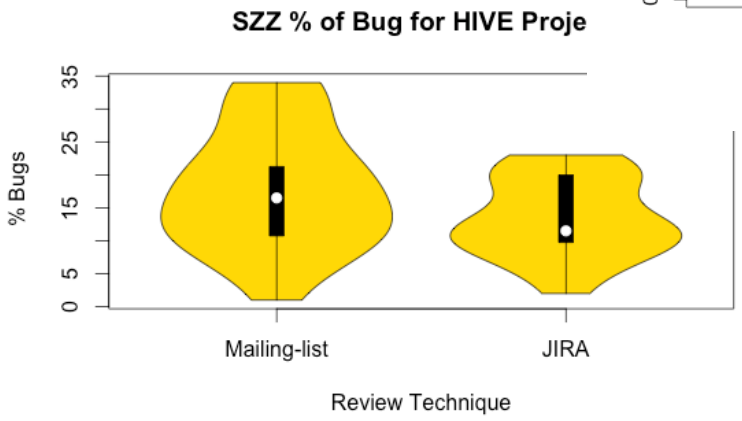

SZZ % of Bug for FLEX Project
% Bugs
0
5
15
25
35
Mailing-list
JIRA
Review Technique

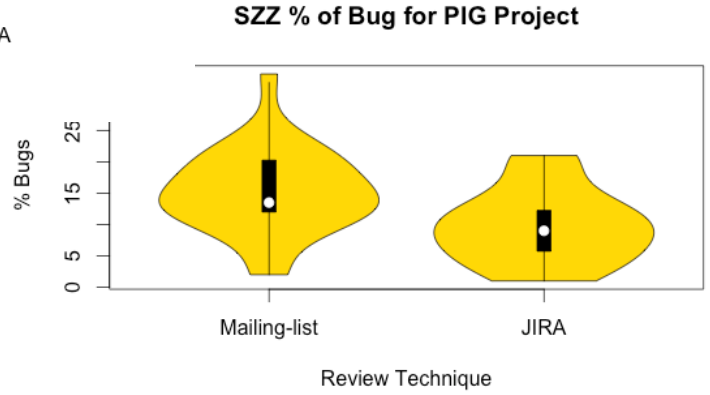
Findings


SZZ % of Bug for FLEX Project
% Bugs
0
5
15
25
35
Mailing-list
JIRA
Review Technique


SZZ % of Bug for HBase Project
% Bugs
20
30
40
Mailing-list
JIRA
Review Technique


SZZ % of Bug for SVN Project
% Bugs
0
10
20
30
40
50
Mailing-list
JIRA
Review Technique


SZZ % of Bug for HIVE Proje
% Bugs
0
5
15
25
35
Mailing-list
JIRA
Review Technique


SZZ % of Bug for PIG Project
% Bugs
0
5
15
25
Mailing-list
JIRA
Review Technique

RQ2: Is the effectiveness of a patch reviewing process related to the medium used?

- **Reviews performed on unicast technology are more effective in terms of catching.**

RQ3: Is the efficiency of a patch review process related to the medium used?

- ❑ Review length (in days) (RL)
- ❑ Response delay (RD)

Approach

- **H0 :** ***There is no significant difference between the value of metric m for patches reviewed on broadcasts and those reviewed on unicast***

❑ **Mann-Whitney U test**

❑ **Cliff's Delta effect Size**

Findings

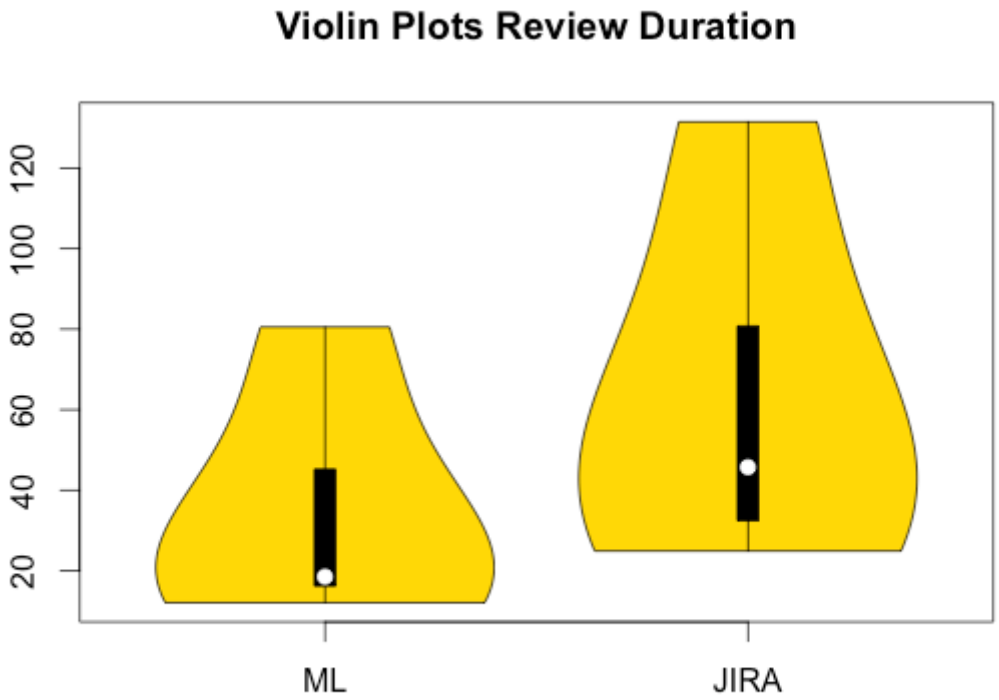
	Review length		Response delay	
	P-values	Cliff's d	P-values	Cliff's d
Flex	0.015	Large	0.019	Large
HBase	0.013	Large	0.016	Large
Hive	0.021	Large	0.011	Large
Pig	0.015	Large	0.012	Large
Subversion	0.017	Large	0.022	Large

Findings

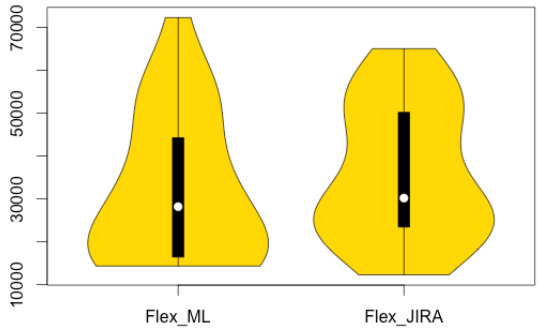
	Review length		Response delay	
	P-values	Cliff's d	P-values	Cliff's d
Flex	0.015	Large	0.019	Large
HBase	0.013	Large	0.016	Large
Hive	0.021	Large	0.011	Large
Pig	0.015	Large	0.012	Large
Subversion	0.017	Large	0.022	Large

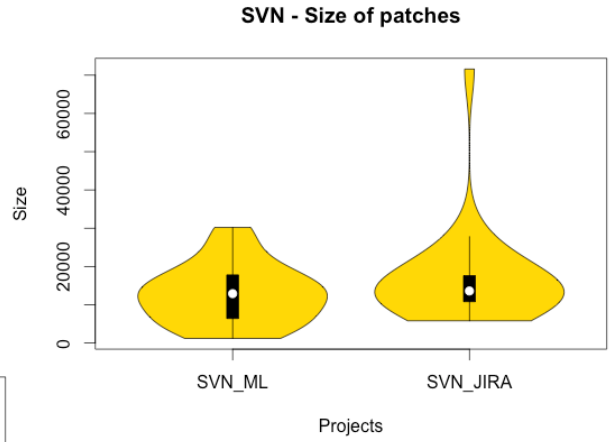
The review length and the response delay is shorter on broadcast

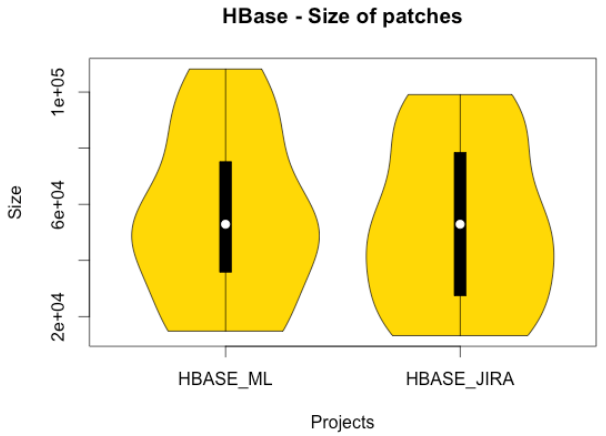
Findings

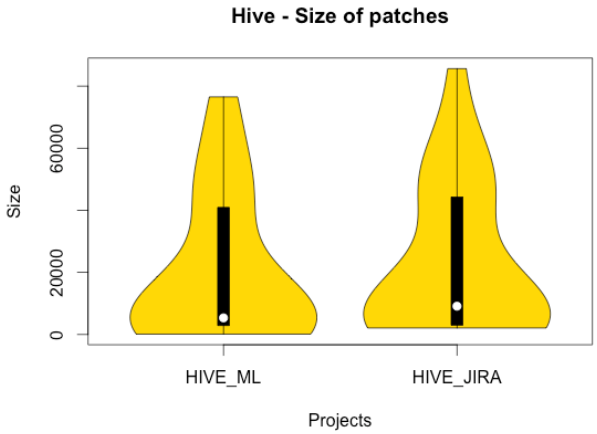


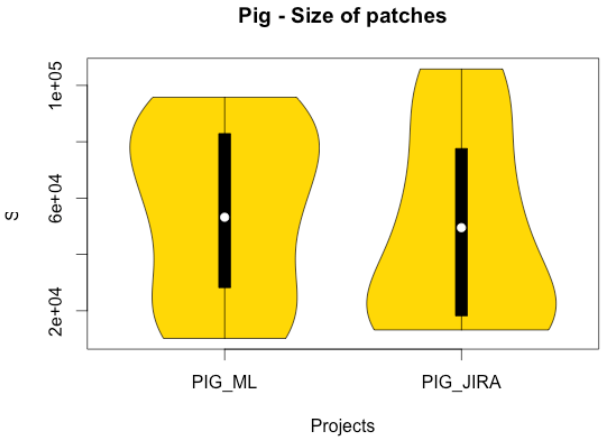
Size of the patches











RQ3: Is the efficiency of a patch review process related to the medium used?

✓YES

- **Broadcast has a short response delay and a shorter review length.**

Survey of Developers

20 participants

Survey of Developers

- **Q: What motivated the switch from broadcast to unicast?**

- **A:** *The broadcast **is good for discussion** (functional/design/release etc.).*

- **A:** *Unicast technology **makes it easier to review patches, track progress on bugs/issues, look up details on old issues**, easier to make release notes on what has been fixed, and easier to organize releases.*

Survey of Developers

- "New developers **learn about the code structure faster with broadcast** than using unicast."

- "The **traffic of patches circulating on broadcast is high**, because it circulates among all those who are subscribed to the broadcast medium."

Limitations

- SZZ heuristic
- The time window sizes
- 5 Subject Systems

Code review:

What is in that name?

RQ1: Is review effort related to the review medium used?

✓YES

- ❑Patches reviewed on unicast undergo more iterations, and
- ❑Unicast's reviewers are more active during code review.

22

RQ2: Is the effectiveness of a patch reviewing process related to the medium used?

✓YES

- ❑ Reviews performed on unicast technology are more effective in terms of catching.

29

RQ3: the efficiency of a patch review process related to the medium used?

✓YES

- ❑ Broadcast has a short response delay and a shorter review length.

35

More details are available in the paper here:
http://swat.polymtl.ca/~foutsekh/docs/ICST-Tita.pdf