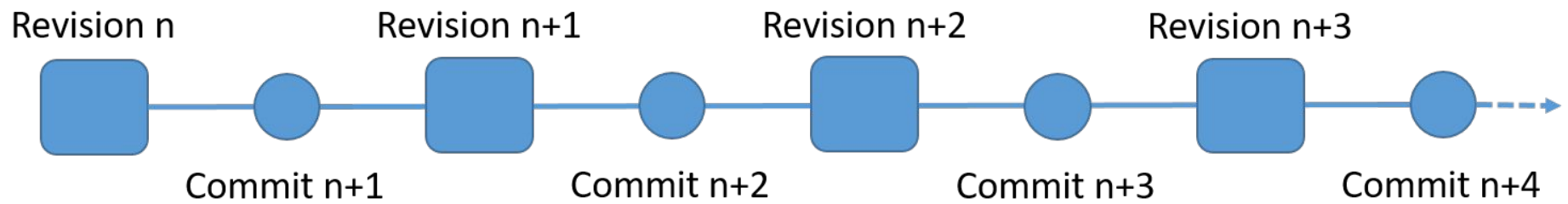


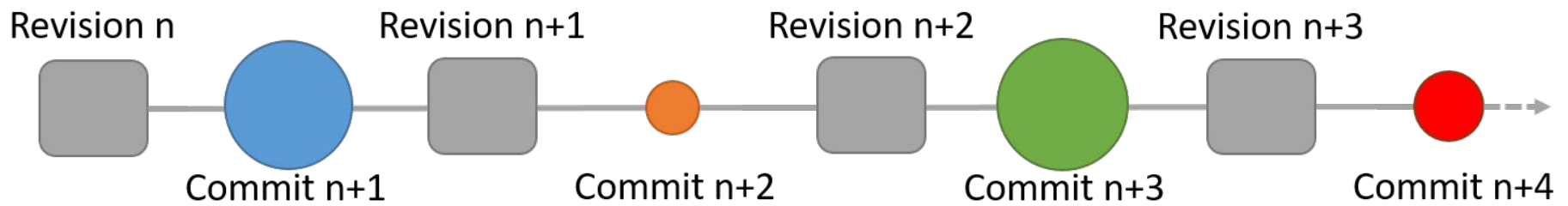
Big Commit Analysis

Andreas Hohler
SCG Seminar Fall 2016
Supervised by Haidar Osman

Software Evolution



Software Evolution



What's a Big Commit?

Example of a small bug fix commit¹): “FLUME-2672 NPE in KafkaSourceCounter”

```
flume-ng-core/src/main/java/org/apache/flume/instrumentation/kafka/KafkaSourceCounter... CHANGED
@@ -31,7 +31,7 @@
 31 31     "source.kafka.empty.count";
 32 32
 33 33     private static final String[] ATTRIBUTES =
 34 -     {TIMER_KAFKA_COMMIT, TIMER_KAFKA_EVENT_GET};
 34 +     {TIMER_KAFKA_COMMIT, TIMER_KAFKA_EVENT_GET, COUNTER_KAFKA_EMPTY};
 35 35
 36 36     public KafkaSourceCounter(String name) {
 37 37         super(name, ATTRIBUTES);
```

1) <https://issues.apache.org/jira/browse/FLUME-2672>

What's a Big Commit?

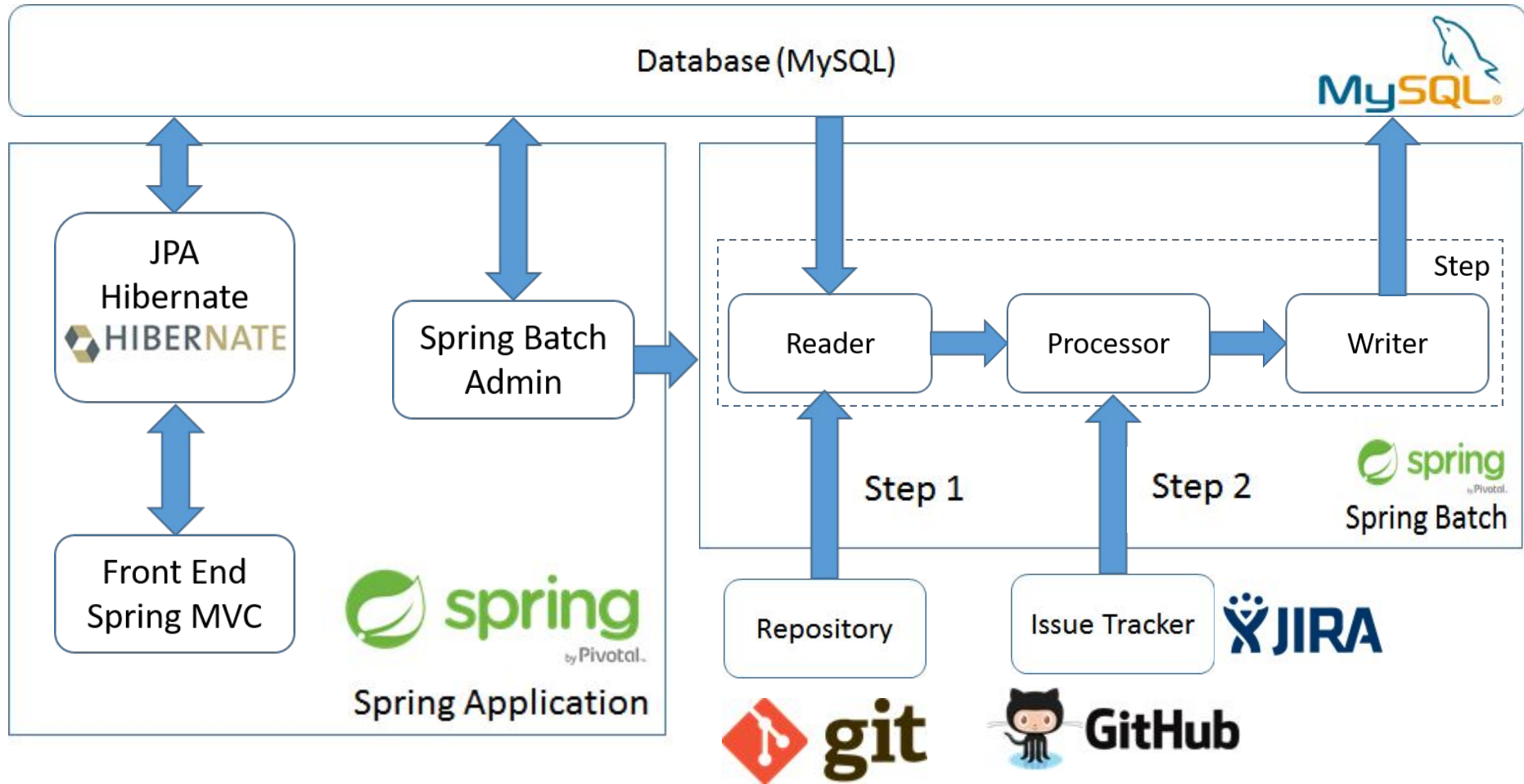
Example of a big commit²⁾ with

- 3'945 LOC changed
- 73 files changed
- LUCENE-6271 issue classified as Bug

Files changed (73) <small>hide</small>		
lucene/core/src/java/org/apache/lucene/search/spans/SpanNotQuery.java	+143	-71
lucene/core/src/java/org/apache/lucene/search/spans/NearSpansUnordered.java	+182	-265
lucene/core/src/test/org/apache/lucene/index/TestIndexWriterDeleteByQuery.java	+71	-0
lucene/core/src/java/org/apache/lucene/search/spans/NearSpans.java	+103	-0
lucene/core/src/java/org/apache/lucene/search/spans/NearSpansPayloadOrdered.java	+146	-0
solr/bin/solr.in.sh	+0	-1
lucene/misc/src/java/org/apache/lucene/uninverting/DocTermOrds.java	+1	-1
solr/solrj/src/java/org/apache/solr/common/params/CommonParams.java	+4	-0
solr/core/src/java/org/apache/solr/core/SolrCore.java	+9	-8
lucene/core/src/test/org/apache/lucene/search/spans/TestBasics.java	+0	-41
lucene/core/src/test/org/apache/lucene/search/spans/JustCompileSearchSpans.java	+42	-33
solr/core/src/java/org/apache/solr/spelling/suggest/FileDictionaryFactory.java	+1	-1
lucene/core/src/test/org/apache/lucene/search/spans/TestFieldMaskingSpanQuery.java	+29	-57
solr/core/src/java/org/apache/solr/spelling/suggest/fst/AnalyzingInfixLookupFactory.java	+1	-1
solr/core/src/java/org/apache/solr/core/RequestParams.java	+1	-0
lucene/facet/src/java/org/apache/lucene/facet/RandomSamplingFacetsCollector.java	+1	-1
lucene/core/src/java/org/apache/lucene/search/spans/SpanWeight.java	+7	-7
lucene/core/src/java/org/apache/lucene/search/payloads/PayloadSpanUtil.java	+10	-6
lucene/highlighter/src/test/org/apache/lucene/search/postingshighlight/TestMultiTermHighlighting...	+1	-1
lucene/core/src/test/org/apache/lucene/search/spans/MultiSpansWrapper.java	+18	-140
lucene/CHANGES.txt	+20	-0
lucene/core/src/java/org/apache/lucene/util/Version.java	+7	-0
lucene/core/src/java/org/apache/lucene/search/ConjunctionDISI.java	+17	-13
solr/core/src/java/org/apache/solr/handler/SolrConfigHandler.java	+264	-12
lucene/core/src/test/org/apache/lucene/search/spans/TestPayloadSpans.java	+67	-70
solr/core/src/java/org/apache/solr/handler/admin/LoggingHandler.java	+1	-1

2) <https://issues.apache.org/jira/browse/LUCENE-6271>

Tool Architecture



Dataset

Projects from GitHub



elastic

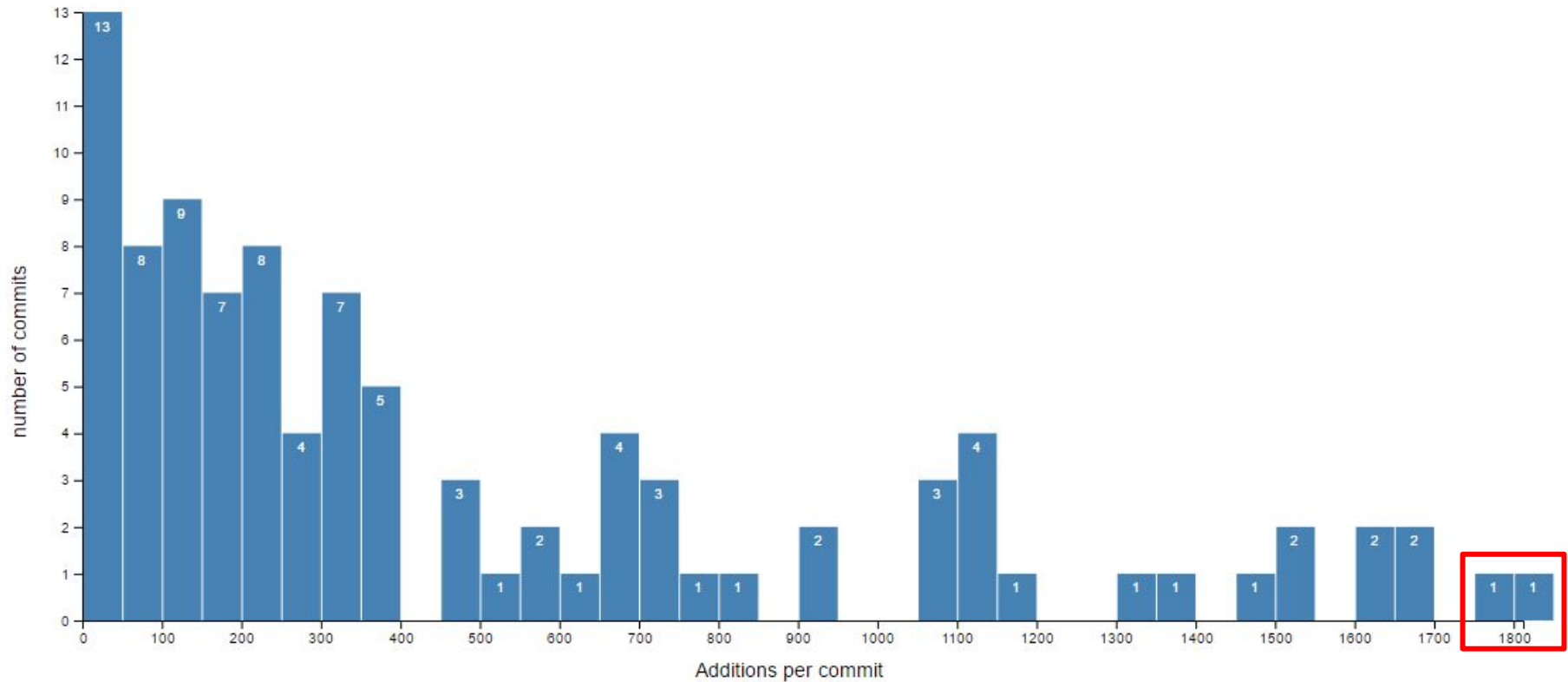


Demo

Analysis

Project: Flume, Issue type: Feature

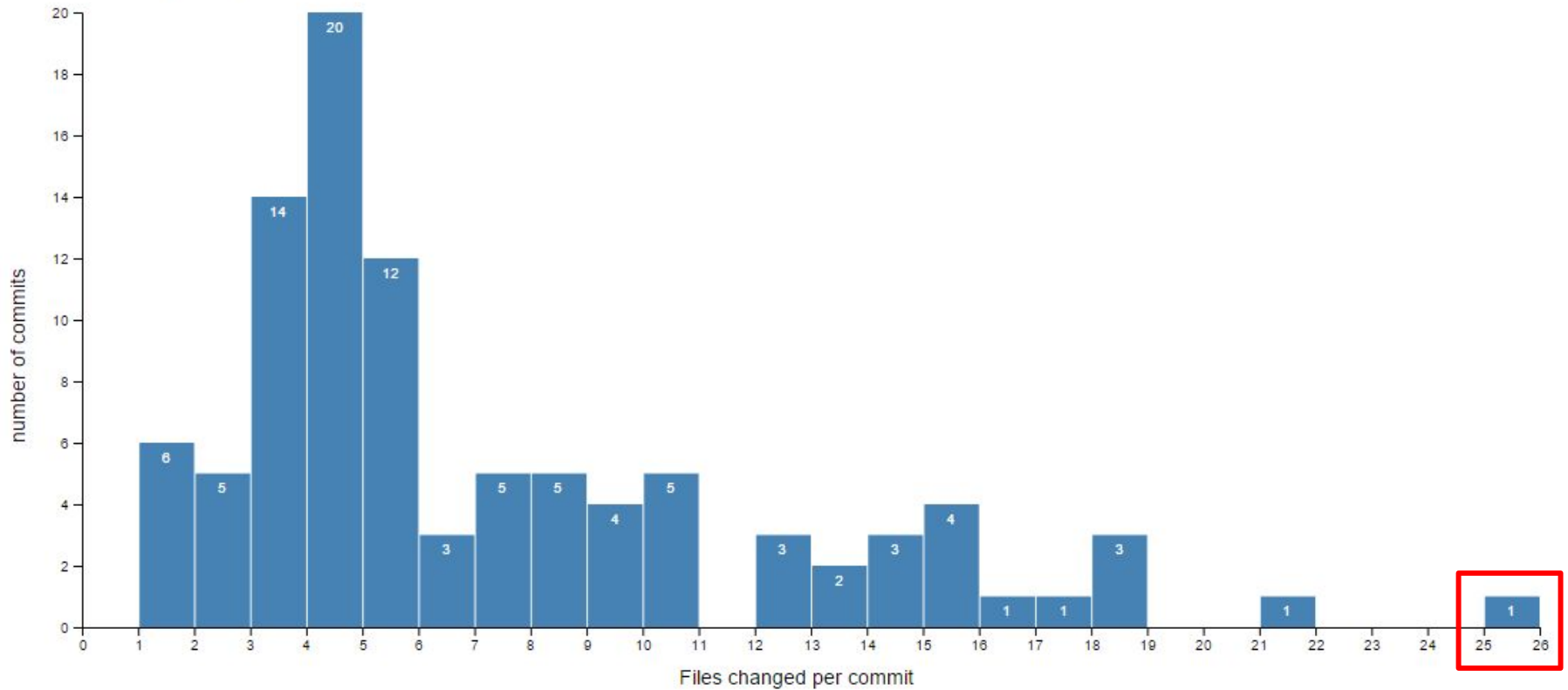
Additions per commit



Analysis

Project: Flume, Issue type: Feature

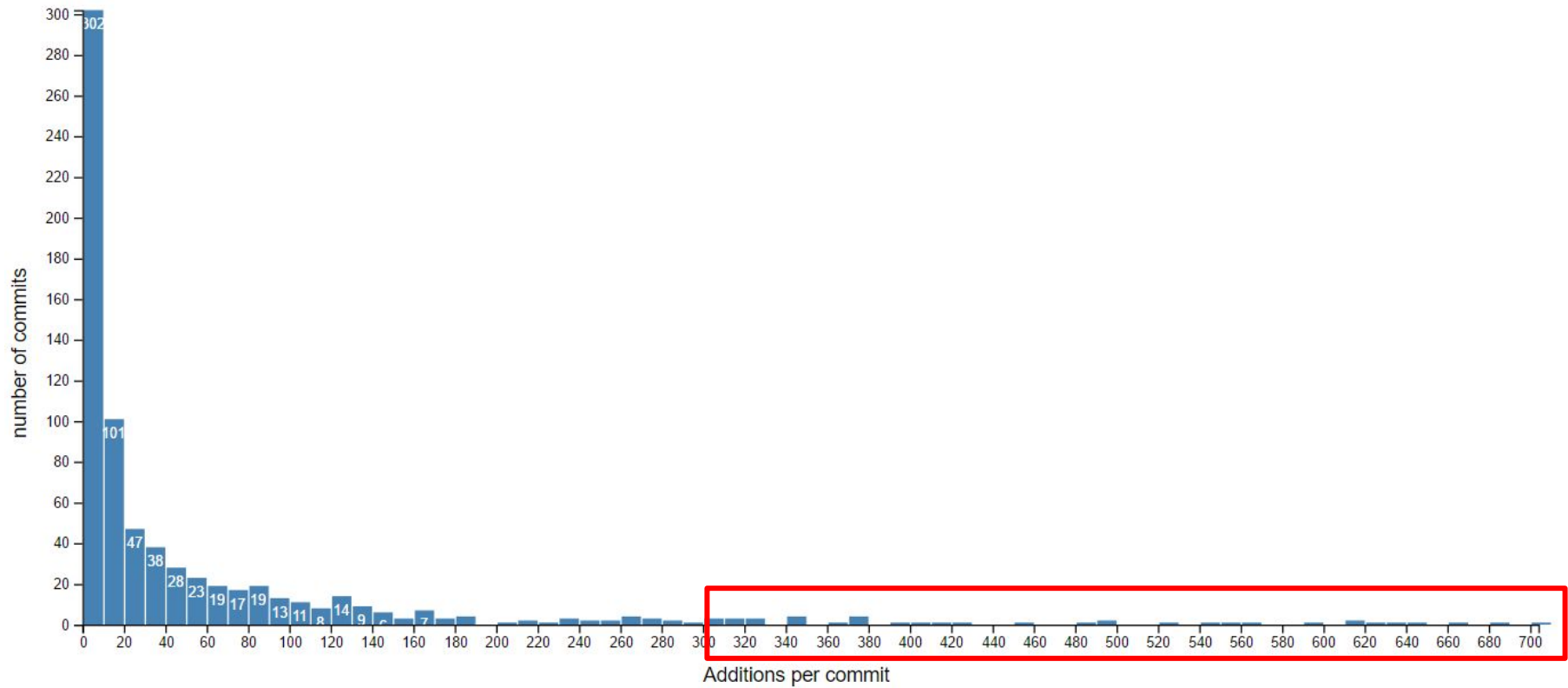
Files changed per commit



Analysis

Project: Flume, Issue type: Bug

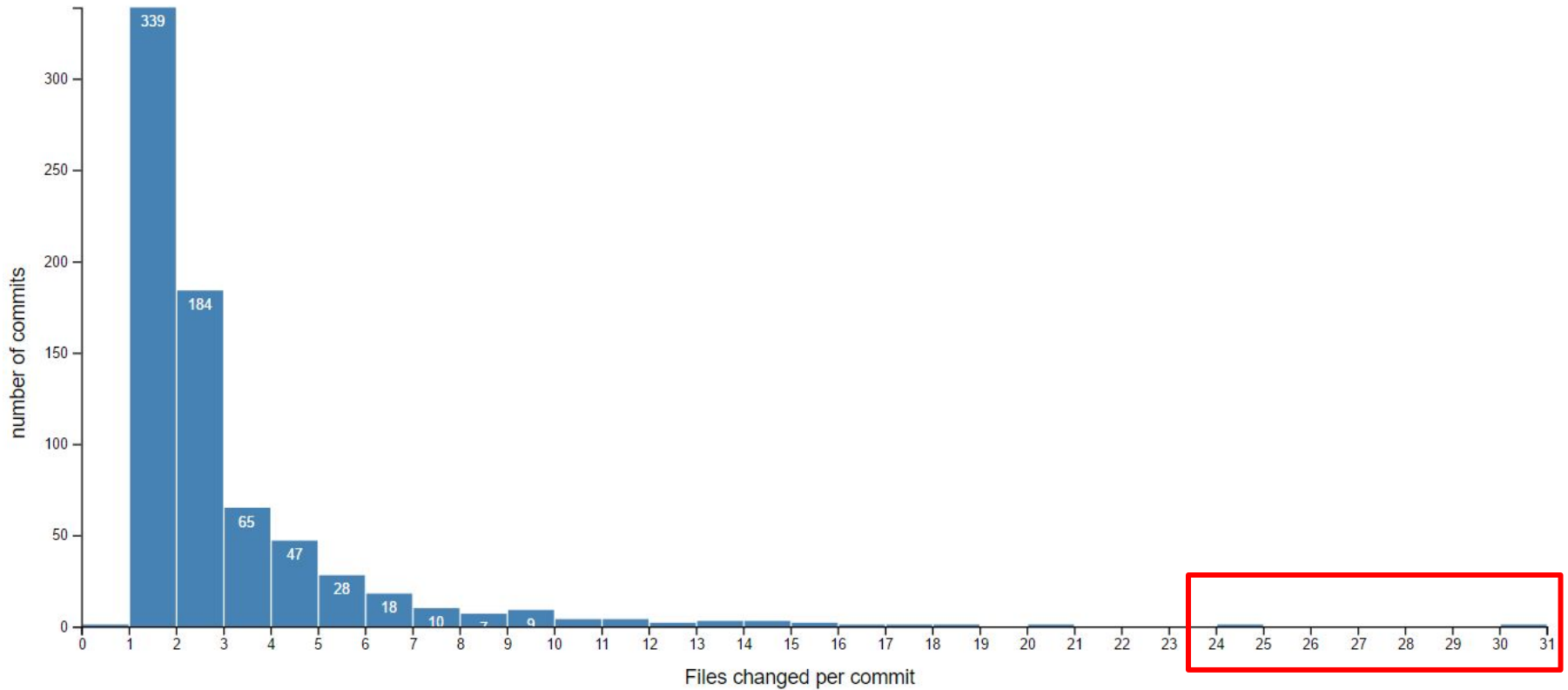
Additions per commit



Analysis

Project: Flume, Issue type: Bug

Files changed per commit



ID	1
Name	flume

Issue types

[Bug](#)
[Feature](#)
[Improvement](#)
[Refactor](#)
[Documentation](#)

Feature (108 results)

Number of issues	108	Highest additions per commit	14122
Median files changed (commits)	5	Lowest additions per commit	3
Median additions (commits)	322	Highest deletions per commit	1596
Median deletions (commits)	8	Lowest deletions per commit	0
Median additions (files)	11	Highest additions per file	1828
Median deletions (files)	0	Lowest additions per file	0
Highest Number of files changed per commit	323	Highest deletions per file	1048
Lowest Number of files changed per commit	1	Lowest deletions per file	0

Additions per commit

