

ISW2 Project A.A. 2019-2020

Andrea Graziani (0273395)
andrea.graziani93@outlook.it
Università degli Studi di Roma "Tor Vergata"
Rome, Italy

KEYWORDS

Testing, Equivalence Class Partitioning, Boundary Value Analysis, Mutation Testing

1 INTRODUCTION

In this report we will describe all results and issues about the utilization of several *testing techniques* involving following open source projects:

Apache BookKeeper™ Which is a scalable, fault tolerant and low latency storage service¹.

Apache OpenJPA™ An implementation of the Java Persistence API specification².

To be more precise, the aim of this report is to analyse 2 classes of each aforementioned project, attempting both to find if there are any errors in it and to increase our confidence in the correct functioning of the **software under test (SUT)**, although the completeness of our testing does not necessarily demonstrate that these classes are error free.

1.1 Class selection

To select classes for our testing activities, we have mainly relied on results elaborated by our **defect prediction model**, developed using **Weka**³ machine learning software, using a *random forests* classifier without *sampling* or *feature selection*.

Particularly, we have focused on classes marked as *defective* (or *buggy*) by our predictor because it is likely that they exhibit a defect observable by our test sets. As known, this approach is able to *reduce the cost of testing*, by letting to focus on specific classes, ignoring stable ones [5].

In particular, for our testing activities involving Apache BookKeeper™, we have choose following classes:

DefaultEnsemblePlacementPolicy⁴ Because it has been marked as defective by our prediction model and, as we will see later, our testing activity has confirmed that prediction.

DiskChecker⁵ Although it has not been marked as defective, we have choose this class due to its relative high LOC value (294), high number of revisions (11) and high number of authors (9). Despite its presumed stability, our testing activity has found several defects inside that class.

1.2 Testing technique

In order to build our test set, we have adopted a testing technique called *equivalence class partitioning*, according to which the domain

of possible input data for each input data element is divided into **equivalence classes**; an equivalence class is a set of data values that the tester assumes are processed in the same way by the test object [8].

Equivalence class definition was been mainly based on *specifications* of our test objects using, therefore, a *black-box* approach; however, sometimes, due to lacking of detailed specifications, we have performed our analysis looking the code too, using conversely a **while-box** approach.

For any equivalence classes is important to find a **representative**. As known, testing one representative of the equivalence class is considered sufficient because it is assumed that for any other input value of the same equivalence class, the test object will show the same reaction or behaviour [8]. Besides equivalence classes for *correct* input, those for *incorrect* input values must be tested as well.

To guarantee that all test object reactions are triggered, we have combined representative values using following rules:

- (1) The representative value of all valid equivalence classes have been combined to test cases (*valid/positive test case*), meaning that all possible combinations of valid equivalence classes will be covered.
- (2) The representative value of an invalid equivalence class have been combined only with representatives of other valid equivalence classes (*invalid/negative test case*).

Moreover, we have adopted following guidelines[8]:

- (1) Test cases including boundary values combinations are preferred.
- (2) Every representative of an equivalence class appears in at least one test case.
- (3) Representatives of invalid equivalence classes should not be combined with representatives of other invalid equivalence classes.

2 APACHE BOOKKEEPER™

2.1 DefaultEnsemblePlacementPolicy class testing analysis

What are responsibilities of DefaultEnsemblePlacementPolicy class? What meant by ensemble? And what for placement policy?

According to bookkeeper specifications[1], an **ensemble** represents a group of **bookies** storing **entries**. To be more precise, according to bookkeeper's nomenclature, a **bookie** is an individual storage server while **entries** represent stored data, therefore an ensemble of size E represents simply a group of storage servers.

The aim of this design is to guarantee **consistency** in an ensemble of bookies of all stored data exploiting a **quorum-based replicated-write** protocol. As known, to support replicated writes at multiple replicas of a file, a client must first contact at least half the servers plus one (a majority) and get them to agree to do the

¹<https://github.com/apache/bookkeeper>

²<https://github.com/apache/openjpa>

³<https://www.cs.waikato.ac.nz/~ml/weka/>

⁴`org.apache.bookkeeper.client.DefaultEnsemblePlacementPolicy`

⁵`org.apache.bookkeeper.util.DiskChecker`

update[9]. Technically, in order to modify a file, a client needs to assemble the so called **write quorum**, that is an arbitrary collection of servers which must be more than the half of all available servers[9]. Therefore, using bookkeeper’s nomenclature, the size of write quorum Q_w represents the number of bookies where each entry is written.

According to bookKeeper protocol[3], the following invariant must hold:

$$E \geq Q_w \geq Q_a \quad (1)$$

In other word, the ensemble size E must be larger than the write quorum size Q_w , which must in turn be larger than the so called **ack quorum size** Q_a , which represents, instead, the number of nodes an entry must be acknowledged on.

BookKeeper uses several algorithms to select a number of bookies from a cluster to build an ensemble compliant to above specifications, some of which are capable to exploit several network topology proprieties too.

According to BookKeeper design, the implementations of these algorithm must be compliant to EnsemblePlacementPolicy interface [4]; in other words, any implementation must respect a specific *contract*, established by aforementioned interface, which covers aspects related to initialization and bookie selection for data placement and reads[4].

Currently there are 3 implementations available by default. They are:

- DefaultEnsemblePlacementPolicy
- RackawareEnsemblePlacementPolicy
- RegionAwareEnsemblePlacementPolicy

In particular, DefaultEnsemblePlacementPolicy class encapsulates the simplest algorithm for bookie selection for ensemble creation because it, simply, picks bookies randomly in order to build an ensemble.

2.1.1 Methods testing analysis.

2.1.1.1 initialize.

According to documentation, after the creation of an DefaultEnsemblePlacementPolicy object’s instance, bookKeeper client have to call initialize method, which signature is reported in **listing 1**, in order to initialize the placement policy.

Observing method’s signature and documentation, we note that several complex resources, passed as parameter, are necessary to invoke this method, however, observing method’s implementation, only one of them is effectively used, simplifying our testing activity.

In fact, only conf parameter, representing a ClientConfiguration object, used to contain client configuration necessary for bookKeeper client construction, is effectively used.

Moreover, only following methods, belonging ClientConfiguration class, using conf parameter, are invoked in initialize method:

- getDiskWeightBasedPlacementEnabled
- getBookieMaxWeightMultipleForWeightBasedPlacement

Therefore, in order to test initialize method, we have choose to use a **mock**. As known, a mock is a special-purpose replacement class that mimic behaviour of real objects in controlled ways, for example returning values to the calling component. As known, the

use of mock is very useful during testing activity of incomplete portions of software because you have not to wait for their completion to start your test activity.

Oblivious, we have use **Mockito framework**⁶ to manage conf parameter’s return values.

All valid and invalid equivalence classes, indicated as vEC and iEC respectively, are reported in **table 1**. We have developed 2 valid test cases and 1 negative test case. All test passed.

Listing 1: Signature of method initialize

```
EnsemblePlacementPolicy initialize(ClientConfiguration conf,
    Optional<DNSToSwitchMapping> optionalDnsResolver,
    HashedWheelTimer hashedWheelTimer,
    FeatureProvider featureProvider,
    StatsLogger statsLogger)
```

2.1.1.2 newEnsemble.

Documentation [4] reports that newEnsemble method is used to build an ensemble made up of several bookies. Method signature, shown below in **listing 2**, reports that it takes up to 5 parameters and return an ensemble which is represented by a List<BookieSocketAddress> object.

Listing 2: Signature of method newEnsemble

```
List<BookieSocketAddress> newEnsemble(int ensembleSize,
    int writeQuorumSize,
    int ackQuorumSize,
    Map<String,byte[]> customMetadata,
    Set<BookieSocketAddress> excludeBookies)
    throws BKException.BKNotEnoughBookiesException
```

The meaning of all parameter is:

ensembleSize represents the ensemble size.

writeQuorumSize the value of write quorum size.

ackQuorumSize the value of ack quorum Size.

customMetadata user meta-data.

excludeBookies a collection of bookies that should not be considered as targets for the new ensemble.

Is very important to precise that, in order to build an ensemble, bookies are picked up from a set called knownBookies, which can be populated invoking onClusterChanged method (which we will describe following).

Moreover, please note that we will use K to represent the set of BookieSocketAddress objects of knownBookies, while E , conversely, will be used to refer to the set of BookieSocketAddress objects belonging to excludeBookies.

All valid and invalid equivalence classes are reported in **table 2**.

For this method, we have developed $2 \times 1 \times 1 \times 1 \times 2 = 4$ valid test cases, by combining the representatives of the equivalence classes, and $2 + 1 + 1 + 3 = 7$ negative test cases, by separately testing representatives of every invalid class, with a total of 11 test cases. However, every test case was been executed twice because, as said previously, two different forms of initialization of DefaultEnsemblePlacementPolicy object exist (see initialize method); therefore, we have executed a total of 22 tests; of these, 8 failed.

⁶<https://site.mockito.org/>

Table 1: Equivalence classes and representatives of initialize method

| Parameter | Equivalence Classes | Representatives |
|---------------------|--|-----------------|
| conf | vEC_1 A not-null ClientConfiguration object which: • <code>conf.getDiskWeightBasedPlacementEnabled() = true</code> | (see the code) |
| | vEC_1 A not-null ClientConfiguration object which: • <code>conf.getDiskWeightBasedPlacementEnabled() = false</code> | (see the code) |
| | iEC_1 null object | null |
| optionalDnsResolver | vEC_1 Any value. optionalDnsResolver is never used or accessed by the class. | null |
| hashedWheelTimer | vEC_1 Any value. hashedWheelTimer is never used or accessed by the class. | null |
| featureProvider | vEC_1 Any value. featureProvider is never used or accessed by the class. | null |
| statsLogger | vEC_1 Any value. statsLogger is never used or accessed by the class. | null |

Table 2: Equivalence classes and representatives of newEnsemble method

| Parameter | Equivalence Classes | Representatives |
|-----------------|---|-------------------|
| ensembleSize | vEC_1 ensembleSize = 0 | 0 |
| | vEC_2 $0 < \text{ensembleSize} \leq K - E $ | $ K - E $ |
| | iEC_1 ensembleSize < 0 | -1 |
| | iEC_2 ensembleSize > $ K - E $ | $ K - E + 1$ |
| writeQuorumSize | vEC_1 writeQuorumSize ≤ ensembleSize | ensembleSize |
| | iEC_1 writeQuorumSize > ensembleSize | ensembleSize + 1 |
| ackQuorumSize | vEC_1 ackQuorumSize ≤ writeQuorumSize | writeQuorumSize |
| | iEC_1 ackQuorumSize > writeQuorumSize | ackQuorumSize + 1 |
| customMetadata | vEC_1 Any value. customMetadata is never used or accessed by the class. | null |
| excludeBookies | vEC_1 A not-null Set<BookieSocketAddress> object where: • $ E = 0$ | new HashSet<>() |
| | vEC_2 A not-null Set<BookieSocketAddress> object where: • $0 < E < K $ • $E \subset K$ • $ K - E \geq \text{ensembleSize}$ | (see the code) |
| | iEC_1 A not-null Set<BookieSocketAddress> object where: • $ E > 0$ • $\exists x \in E : x \notin K$ | (see the code) |
| | iEC_2 A not-null Set<BookieSocketAddress> object where: • $ E > 0$ • $\exists x \in E : x = \text{null}$ | (see the code) |
| | iEC_3 null object | null |

2.1.1.3 onClusterChanged.

onClusterChanged method is used to update the view of the cluster, that is to specify what bookies are available as *writable* and what bookies are available as *read-only*; this operation is necessary to populate, or update, knownBookies set which, as already said, is used to pick up bookies during newEnsemble invocation. According to documentation, onClusterChanged should be invoked when any changes happen in the cluster, returning a list of *failed* (or *dead*) bookies during this cluster change.

Signature is reported in **listing 3**.

Listing 3: Signature of method onClusterChanged

```
Set<BookieSocketAddress> onClusterChanged(Set<BookieSocketAddress> writableBookies,
                                         Set<BookieSocketAddress> readOnlyBookies)
```

All valid and invalid equivalence classes are reported in **table 3**.

Combining the representatives of the equivalence classes, we have developed are $2 \times 2 = 4$ valid test cases and $3 + 3 = 6$ negative test cases, by separately testing representatives of every invalid class. Therefore, we have a 10 distinct tests and, since we have to

run each test twice, we have a total of 20 test cases; of these, 8 failed.

2.1.1.4 updateBookieInfo.

According to DefaultEnsemblePlacementPolicy interface's documentation [4], updateBookieInfo is used to update bookie info details, taking only one input parameter, a Map<BookieSocketAddress, BookieInfo> object.

Please note that we will use K to represent the set of the *keys* (or *indexes*) of the dictionary bookieInfoMap, that is the set of all BookieSocketAddress objects used as indexes. Conversely, V will be used to refer to the set of the *values* of the dictionary, made up of BookieInfo objects.

Signature is reported in **listing 4**, while equivalence classes are shown in **table 4**. In this case, we have developed are 2 valid test cases and 4 negative test cases. Therefore, we have 6 distinct tests and, since we have to run each test twice, we have a total of 12 test cases; of these, 3 failed.

Listing 4: Signature of method updateBookieInfo

Table 3: Equivalence classes and representatives of onClusterChanged method

| Parameter | Equivalence Classes | Representatives |
|-----------------|--|-----------------|
| writableBookies | vEC_1 An not-null Set<BookieSocketAddress> object where: • writableBookies.size() = 0 | new HashSet<>() |
| | vEC_2 A not-null Set<BookieSocketAddress> object where: • writableBookies.size() > 0 • writableBookies ∩ readOnlyBookies = ∅ | (see the code) |
| | iEC_1 A not-null Set<BookieSocketAddress> object where: • writableBookies.size() > 0 • $\exists x \in \text{writableBookies} : x \in \text{readOnlyBookies}$ | (see the code) |
| | iEC_2 A not-null Set<BookieSocketAddress> object where: • writableBookies.size() > 0 • $\exists x \in \text{writableBookies} : x = \text{null}$ | (see the code) |
| | iEC_3 null object | null |
| readOnlyBookies | vEC_1 An not-null Set<BookieSocketAddress> object where: • readOnlyBookies.size() = 0 | new HashSet<>() |
| | vEC_2 A not-null Set<BookieSocketAddress> object where: • readOnlyBookies.size() > 0 • writableBookies ∩ readOnlyBookies = ∅ | (see the code) |
| | iEC_1 A not-null Set<BookieSocketAddress> object where: • readOnlyBookies.size() > 0 • $\exists x \in \text{readOnlyBookies} : x \in \text{writableBookies}$ | (see the code) |
| | iEC_2 A not-null Set<BookieSocketAddress> object where: • readOnlyBookies.size() > 0 • $\exists x \in \text{readOnlyBookies} : x = \text{null}$ | (see the code) |
| | iEC_3 null object | null |

```
void updateBookieInfo(Map<BookieSocketAddress, BookieInfo> bookieInfoMap)
```

2.1.1.5 replaceBookie.

Documentation reports that replaceBookie is choose randomly a bookie from knownBookies set, if available, in order to replace it with a bookie passed as parameter.

Observing method's implementation and its signature, the latter reported in **listing 5**, it is easy to notice a similarity with newEnsemble the method, described previously. Despite very little differences, equivalence classes of both methods are, practically, the same; please see **table 5**.

Listing 5: Signature of method replaceBookie

```
BookieSocketAddress replaceBookie(int ensembleSize,
    int writeQuorumSize,
    int ackQuorumSize,
    Map<String, byte[]> customMetadata,
    Set<BookieSocketAddress> currentEnsemble,
    BookieSocketAddress bookieToReplace,
    Set<BookieSocketAddress> excludeBookies)
    throws BKException, BKNotEnoughBookiesException
```

2.1.1.6 Other methods.

To be precise, DefaultEnsemblePlacementPolicy class contains many other methods not analysed in this report because either they does nothing or they return always the same value regardless input parameters.

Those methods are:

unitalize This method does nothing.

registerSlowBookie This method does nothing too.

isEnsembleAdheringToPlacementPolicy Return always PlacementPolicyAdherence.MEETS_STRICT regardless input parameters.

reorderReadSequence Returns always a WriteSet object which is passed as parameter.

Oblivious, for purposes related to our adequacy criteria, we have developed one test case for each above method, passing random values as parameters.

2.1.2 Adequacy Criteria.

Is our test set T good enough? Has DefaultEnsemblePlacementPolicy class been tested thoroughly?

In general, test adequacy refers to the *goodness* of a test set, which must be measured against a quantitative criterion. As known, a test adequacy criterion could be based either on *requirements* or the *implementation* of the program under test. Particularly, for all test objects analysed in this report, we have adopted several **white-box** test adequacy criteria, that is depend solely on the implementation of our test objects.

2.1.2.1 Statement Coverage.

The statement coverage of a test set T is computed as following[7]:

$$\text{Statement Coverage} = \frac{|S_c|}{|S_e| - |S_i|} \quad (2)$$

Where:

S_c is the set of statements covered.

S_i is the set of unreachable statements.

Table 4: Equivalence classes and representatives of updateBookieInfo method

| Parameter | Equivalence Classes | Representatives |
|---------------|---|-----------------|
| bookieInfoMap | vEC_1 : A not-null Map<BookieSocketAddress, BookieInfo> object where: • <code>bookieInfoMap.size() = 0</code> | (see the code) |
| | vEC_2 : A not-null Map<BookieSocketAddress, BookieInfo> object where: • <code>bookieInfoMap.size() > 0</code> • $\forall (x, y) \in K \times V \Rightarrow x \neq null, y \neq null$ | (see the code) |
| | iEC_1 : A not-null Map<BookieSocketAddress, BookieInfo> object where: • <code>bookieInfoMap.size() > 0</code> • $\exists (x, y) \in K \times V : x = null, y \neq null$ | (see the code) |
| | iEC_2 : A not-null Map<BookieSocketAddress, BookieInfo> object where: • <code>bookieInfoMap.size() > 0</code> • $\exists (x, y) \in K \times V : x \neq null, y = null$ | (see the code) |
| | iEC_3 : A not-null Map<BookieSocketAddress, BookieInfo> object where: • <code>bookieInfoMap.size() > 0</code> • $\exists (x, y) \in K \times V : x = null, y = null$ | (see the code) |
| | iEC_4 : null object. | null |

Table 5: Equivalence classes and representatives of replaceBookie method

| Parameter | Equivalence Classes | Representatives |
|-----------------|--|-------------------|
| ensembleSize | See table 2 | |
| writeQuorumSize | See table 2 | |
| ackQuorumSize | See table 2 | |
| customMetadata | See table 2 | |
| currentEnsemble | vEC_1 : A not-null List<BookieSocketAddress> object where: • <code>currentEnsemble.size() > 0</code> • $\forall x \in currentEnsemble \Rightarrow x \in K$ • $\forall x \in currentEnsemble \Rightarrow x \neq null$ | (see the code) |
| | iEC_1 : A not-null List<BookieSocketAddress> object where: • <code>currentEnsemble.size() = 0</code> | new ArrayList<>() |
| | iEC_2 : A not-null List<BookieSocketAddress> object where: • <code>currentEnsemble.size() > 0</code> • $\forall x \in currentEnsemble \Rightarrow x \neq null$ • $\exists x \in currentEnsemble : x \notin K$ | (see the code) |
| | iEC_3 : A not-null List<BookieSocketAddress> object where: • <code>currentEnsemble.size() > 0</code> • $\exists x \in currentEnsemble : x = null$ | (see the code) |
| | iEC_4 : null object | null |
| bookieToReplace | vEC_1 : A valid BookieSocketAddress object where: • <code>bookieToReplace \in currentEnsemble</code> • <code>bookieToReplace \neq null</code> | (see the code) |
| | iEC_1 : A valid BookieSocketAddress object where: • <code>bookieToReplace \notin currentEnsemble</code> • <code>bookieToReplace \neq null</code> | (see the code) |
| | iEC_2 : null object | null |
| excludeBookies | See table 2 | |

S_e is the set of statements in the program.

Let's start with our test set T_1 , that is the test set containing all valid and invalid test cases built during the analysis of the equivalence classes made previously.

According to sonarcloud and JaCoCo reports, our test set T_1 is able to cover up to 72 lines statements out of a total of 82, giving to us a statement coverage equal to 0.963 (96.3%)

To improve statement coverage, we have add following unit test, obtaining test set T_2 :

- `additionalTestCase_1()` (TestOnClusterChanged class)

Fortunately, test set T_2 is able to cover up to 81 lines statements out of a total of 82.

However, we are unable to cover all statements because, we believe, in `newEnsemble` method is present an *unreachable* statements,

because it falls on an *infeasible path*, that is a path that would never be reached by any test set with any type of input data.

Therefore, since $|S_i| = 1$, according to **eq. (2)**, statement coverage is equal to 1 (100%).

Since statement coverage of T_2 is 1, we can consider T_2 as **adequate with respect to the statement coverage criterion**.

2.1.2.2 Decision coverage (Branch decision coverage).

According to [7], a *decision* is considered covered if the flow of control has been diverted to all possible destinations that correspond to this decision, i.e. all outcomes of the decision have been taken.

This implies that, for example, the expression in the if or while statement has evaluated to true in some execution and to false in the same or another execution. Note that each if and each while contribute to *one* decision whereas a switch may contribute to more than one [7].

Decision coverage of a test set T is computed as following[7]:

$$\text{Decision Coverage} = \frac{|D_c|}{|D_e| - |D_i|} \quad (3)$$

Where:

D_c is the set of decisions covered.

D_i is the set of unreachable decisions.

D_e is the set of decisions in the program.

According to our analysis, our test set T_2 is able to cover 14 decisions out of a total of 15. In newEnsemble method, there is an unreachable decisions, therefore, since $|D_i| = 1$, our test set T_2 give to us a decision coverage equal to 1 (100%).

Since decision coverage of T_2 is 1, we can consider T_2 as **adequate with respect to the decision coverage criterion**.

2.1.2.3 Condition coverage.

Unlike decision coverage, condition coverage ensures that each simple condition within a *compound* condition has assumed both values true and false[7].

Condition coverage of a test set T is computed as following[7]:

$$\text{Condition Coverage} = \frac{|C_c|}{|C_e| - |C_i|} \quad (4)$$

Where:

C_e the set of simple conditions in the program.

C_c the set of simple conditions covered by our test set T .

C_i the set infeasible simple conditions.

Our test set T_2 is able to cover 33 decisions out of a total of 34. Unfortunately, in newEnsemble method there is an infeasible conditions (a simple condition inside a while statement which is never false), therefore, since $|C_i| = 1$, our test set T_2 give to us a condition coverage equal to 1 (100%).

Since condition coverage of T_2 is 1, we can consider T_2 as **adequate with respect to the condition coverage criterion**.

2.1.3 Mutation Analysis.

So, mutation testing represents a very powerful technique to achieve two goals[6]:

- (1) Evaluating test suite quality.
- (2) Once we have a good test suite, executing its test cases against the original program to find errors.

In mutation analysis, from a program p , a set of faulty programs p_m , called *mutants*, is generated by a few single syntactic changes to the original program p . Assume that a test set T is supplied to the system. In this case each, mutant p_m will then be run against this test set T . If the result of running p_m is different from the result of running p for any test case in T , then the mutant p_m is said to be *killed*, otherwise, it is said to have *survived*[6].

Our objective, is to improve our test set T providing additional test inputs (or improving existing ones) to kill these surviving mutants.

As known, to evaluate the goodness of a test suite, an adequacy criterion, known as the *Mutation Score* (or *Mutation Coverage*), is used.

So, according to PIT⁷ report, mutation coverage of our T_2 test set is equal to 28, out of a total of 41 (mutation score 68.29%). This result means that, after that each mutant is been run against our test set T_2 , some mutants *survived*.

Therefore, to improving our test suite, we have add following unit tests to our test set T_2 :

- additionalTestCase_1() (TestInitialize class)
- additionalTestCase_1() (TestOther class)

Moreover, we have improved following unit test:

- test_1 (TestOther class)
- test_3 (TestOther class)
- test_4 (TestOther class)

After these changes, with our new test set T_3 , we have killed 4 extra mutants, obtaining a mutation coverage equal to 32, out of a total of 41 (mutation score 78.04%).

As known, the goal of mutation analysis is to raise the mutation score to 1, indicating that a test set T is sufficient to detect all the faults denoted by the mutants[6]. However, this aim is out the scope of our project.

2.1.4 Conclusions. Finally, using our tests set T_3 , 23 test, out of a total of 86, failed, revealing the presence of several bugs, confirming results made by defect prediction model.

⁷<https://pitest.org/>

2.2 DiskChecker class testing analysis

According to BookKeeper's documentation, DiskChecker class is used to provide several utility functions for checking disk problems, managing all directories belonging to **ledgers**, which represents the basic unit of storage in BookKeeper[1].

Specifications establishes that, for each ledger directory, is possible to specify a maximum disk space D which can be used and a, so-called, warning threshold D_w for disk usage[2]. In particular, the following invariant must hold:

$$\begin{aligned} 0 < D_w < D \\ D_w \leq D \end{aligned} \quad (5)$$

However, for both thresholds, BookKeeper's specifications establish a default value, which is equal to 0.95.

2.2.1 Methods testing analysis.

2.2.1.1 DiskChecker.

Obliviously, that method represents the constructor of DiskChecker class and takes up to 2 parameters:

threshold previously indicated as D
warnThreshold previously indicated as D_w

All valid and invalid equivalence classes, which definition is clearly based on specifications described before, are reported in **table 6**, while method signature is reported in **listing 6**.

From identified equivalence classes, we have developed 1 valid test case and 3 negative test cases. All tests passed. Test cases are reported in **table 7**.

Listing 6: Signature of method DiskChecker

```
public DiskChecker(float threshold, float warnThreshold)
```

2.2.1.2 checkDir.

checkDir method, taking only one File object as parameter, is used to perform several checks involving a directory, verifying disk usage, write and read permissions or the exceeding of the disk usage threshold.

It returns the disk usage fraction usage, defined as following:

$$\text{usage} = 1 - \frac{\text{usableSpace}}{\text{totalSpace}} \quad (6)$$

All valid and invalid equivalence classes are reported in **table 8**, while method signature is reported in **listing 7**. We have developed 2 valid test cases and 5 negative test cases, with a total 8 test cases, all passed. Test cases are described in **table 9**.

During equivalence class definitions, it was assumed that BookKeeper's user is *not* root user.

Listing 7: Signature of method checkDir

```
public float checkDir(File dir) throws DiskErrorException
```

2.2.1.3 getTotalDiskUsage.

getTotalDiskUsage is used to compute disk usage fraction, already defined in **6**, taking as input a list of directories.

All valid and invalid equivalence classes are reported in **table 10**, while method signature is reported in **listing 8**. We have

developed 1 valid test case and 7 negative test cases, with a total of 8 test cases, of which 2 failed.

Listing 8: Signature of method getTotalDiskUsage

```
public float getTotalDiskUsage(List<File> dirs) throws IOException
```

2.2.1.4 getTotalDiskSpace – getTotalFreeSpace.

Observing both the implementations and the signatures of these last methods, called getTotalDiskSpace and getTotalFreeSpace, is easy to understand their meaning and aim, however, if we look at the code carefully, we will note several defects that must be taken into account during test activities:

- (1) Unexpectedly, methods comments are the same, although they perform different activities; in other words, methods comments "*lie*", because they do not describe what the functions truly do.

The only way to build our test set is to observe method implementations ignoring comments, adopting a pure white-box approach.

This consideration is true for getTotalDiskSpace while, for getTotalFreeSpace method, the comment is just imprecise.

- (2) getTotalFreeSpace method signature "*lies*" too. In fact, although we would expect that the method returns "*total free space*" of a set of directories, observing the implementation, is easy to understand that it returns, instead, the *total usable space* of a set of directories.

Since, usable space and free space have different meaning, this erroneous signature is the reason according to which many of our test cases failed.

An other defect, not strictly related to our testing activities, is that methods implementations are the same (except one line). In other words, this is a case of code duplication.

Although some differences (an empty List<File> object is a valid input), equivalence classes of both these methods are the same of the getTotalDiskUsage method, already described in **10**.

To test getTotalDiskSpace method, we have developed 2 valid test cases and 6 negative test cases, with a total 8 test cases, of which 2 failed. Likewise, we have developed the same number of tests for getTotalFreeSpace method too; however 3 unit tests failed out of a total of 8

2.2.2 Adequacy Criteria.

2.2.2.1 Statement coverage.

Our test set T_1 , made up of all test cases built during equivalence class analysis, is able to cover up to 69 statements out of a total of 89, reaching a statement coverage equal to **0,775 (77,5%)**.

To improve our test set, we add following unit tests to test set T_1 :

- (1) additionalTest_1 (TestDiskChecker class).
- (2) additionalTest_1 (TestCheckDir class).
- (3) additionalTest_2 (TestCheckDir class).
- (4) additionalTest_3 (TestCheckDir class).

Table 6: Equivalence classes and representatives of DiskChecker method

| Parameter | Equivalence Classes | Representatives |
|---------------|--|-----------------------|
| threshold | vEC_1 $0 < warnThreshold \leq threshold < 1$ | 0.5 |
| | iEC_1 $threshold \leq 0$ | 0 |
| | iEC_2 $threshold \geq 1$ | 1 |
| | iEC_3 $0 < threshold < warnThreshold < 1$ | $warnThreshold - 0.1$ |
| warnThreshold | vEC_1 : $warnThreshold \leq threshold$ | threshold |
| | iEC_1 $warnThreshold > threshold$ | $threshold + 0.1$ |

Table 7: Test cases of DiskChecker method

| Test Case | Parameter threshold warnThreshold | | Expected output | Actual output | Passed |
|-----------------------------|--------------------------------------|-----|-----------------|---------------|--------|
| <i>Valid</i> ₁ | 0.5 | 0.5 | No Exception | No Exception | ✓ |
| <i>Invalid</i> ₁ | 0 | 0 | Exception | Exception | ✓ |
| <i>Invalid</i> ₂ | 1 | 1 | Exception | Exception | ✓ |
| <i>Invalid</i> ₃ | 0.5 | 0.6 | Exception | Exception | ✓ |

Table 8: Equivalence classes and representatives of checkDir method

| Parameter | Equivalence Classes | Representatives |
|-----------|--|---|
| dir | vEC_1 A not-null File object, which: <ul style="list-style-type: none"> Represents a valid directory Exists BookKeeper's user has read permissions BookKeeper's user has write permissions | <code>createTempDir("directoryFile", "test")</code> |
| | vEC_2 A not-null File object, which: <ul style="list-style-type: none"> Represents a valid directory Not exists BookKeeper's user has write permissions to make that directory | <code>new File("./makeMultiple/dir/path")</code> |
| | iEC_1 A not-null File object, which: <ul style="list-style-type: none"> It does not represent a directory, that is it can be one of the following: <ol style="list-style-type: none"> regular file symbolic link character device file ... | <code>new File("/dev/zero")</code> |
| | iEC_2 A not-null File object, which: <ul style="list-style-type: none"> Represents a valid directory Not exists BookKeeper's user has not write permissions to make that directory | <code>new File("/root/notMakable")</code> |
| | iEC_3 A not-null File object, which: <ul style="list-style-type: none"> Represents a valid directory Exists BookKeeper's user has not read permissions | <code>new File("/root")</code> |
| | iEC_4 A not-null File object, which: <ul style="list-style-type: none"> Represents a valid directory Exists BookKeeper's user has not write permissions | <code>new File("/home")</code> |
| | iEC_5 A not-null File object, which: <ul style="list-style-type: none"> Does not represent a valid directory (from file system point of view) because contains forbidden characters like: <ol style="list-style-type: none"> For NTFS file system: <code>\ / : * ? < > </code> For Btrfs, ext4, ext3, XFS file systems: <code>NULL /</code> | <code>new File("\u0000") (NULL)</code> |
| | iEC_6 A null object. | null |

Table 9: Test cases of checkDir method

| Test Case | Parameter dir | Expected output | Actual output | Passed |
|-----------------------------|--|-----------------|---------------|--------|
| <i>Valid</i> ₁ | IOUtils.createTempDir("directoryFile", "test") | No Exception | No Exception | ✓ |
| <i>Valid</i> ₂ | new File("./makeMultiple/dir/path") | No Exception | No Exception | ✓ |
| <i>Invalid</i> ₁ | new File("/dev/zero") | Exception | Exception | ✓ |
| <i>Invalid</i> ₂ | new File("/root/notMakable") | Exception | Exception | ✓ |
| <i>Invalid</i> ₃ | new File("/root") | Exception | Exception | ✓ |
| <i>Invalid</i> ₄ | new File("/home") | Exception | Exception | ✓ |
| <i>Invalid</i> ₅ | new File("\u0000") | Exception | Exception | ✓ |
| <i>Invalid</i> ₆ | null | Exception | Exception | ✓ |

Table 10: Equivalence classes and representatives of getTotalDiskUsage method

| Parameter | Equivalence Classes | Representatives |
|-----------|--|-----------------------|
| dirs | A not-null List<File> object, which: <ul style="list-style-type: none"> • dirs.size() > 0 • $\forall x \in \text{dirs}$ is true that: <ul style="list-style-type: none"> – x exists – x represents a valid directory – BookKeeper's user has permissions to read x | (see the code) |
| | A not-null List<File> object, where: <ul style="list-style-type: none"> • dirs.size() = 0 | new new ArrayList<>() |
| | A not-null List<File> object, where: <ul style="list-style-type: none"> • dirs.size() > 0 • $\exists x \in \text{dirs} : x$ is not a directory, that is x can be: <ol style="list-style-type: none"> (1) regular file (2) symbolic link (3) character device file (4) ... | (see the code) |
| | A not-null List<File> object, where: <ul style="list-style-type: none"> • dirs.size() > 0 • $\exists x \in \text{dirs} : x$ not exist | (see the code) |
| | A not-null List<File> object, where: <ul style="list-style-type: none"> • dirs.size() > 0 • $\exists x \in \text{dirs} : x = \text{null}$ | (see the code) |
| | A not-null List<File> object, where: <ul style="list-style-type: none"> • dirs.size() > 0 • $\exists x \in \text{dirs} : \text{BookKeeper's user has not permissions to read } x$ | (see the code) |
| | A not-null List<File> object, where: <ul style="list-style-type: none"> • dirs.size() > 0 • $\exists x \in \text{dirs} : x$ does not represent a valid directory (from file system point of view). | (see the code) |
| | <i>iEC</i> ₇ A null object. | null |

With this new test set, T_2 , we are able to cover up to 85 statements out of a total of 89.

However, DiskChecker class contains a method, called setDiskSpaceThreshold(float, float), which, we believe, contains up to 4 unreachable statements. Generally, in order to identify unreachable statements, drawing the flow-graph of the developed code and finding out the path that would never be reached is required. However, in this case, we can consider as unreachable all statements of setDiskSpaceThreshold(float, float) method owing to following reasons:

- (1) setDiskSpaceThreshold(float, float) method is **never used** in the project, that is it is never called by any other method.
- (2) setDiskSpaceThreshold(float, float) method has **no access modifier** (*package private*) which means, according to Java language specification, that it is only accessible within classes in the same package, therefore that method is not visible by our test set.

Therefore, we can conclude by stating that $|S_i| = 4$ and our test set T_2 gives to us a statement coverage equal to **1 (100%)**.

Since statement coverage of T_2 is 1, we can consider T_2 as **adequate with respect to the statement coverage criterion**.

2.2.2.2 Decision coverage (Branch decision coverage).

According to our analysis, DiskChecker class contains 16 decision. Since our test set T_2 covers all decisions, decision coverage is equal to **1 (100%)**. Because decision coverage of T_2 is 1, we can consider T_2 as **adequate with respect to the decision coverage criterion**.

2.2.2.3 Condition coverage.

According to sonarcloud reports, our test set T_2 covers up to 45 simple conditions, out of a total of 46 with a condition coverage equal to 0.978 (97,8%).

Since condition coverage is less than 1, our test set T_2 is *not* adequate with respect to the condition coverage criterion.

2.2.3 Mutation Analysis.

According to PIT report, mutation coverage of our T_2 test set is equal to 29, out of a total of 46 (mutation score 63.04%).

To build our new test set T_3 , we added following unit tests to T_2 :

- additionalTest_1 (TestTest class)
- additionalTest_2 (TestTest class)
- additionalTest_3 (TestTest class)

After these changes, with our new test set T_3 , we have killed 4 extra mutants, obtaining a mutation coverage equal to 34, out of a total of 41 (mutation score 82.92%).

2.2.4 Conclusions. Finally, using our test set T_3 , 7 tests, out of a total of 50, failed revealing the presence of several bugs, despite its stability according to our prediction model.

3 APACHE OPENJPA™

3.1 SimpleRegex class testing analysis

As the name suggests, SimpleRegex class is used to manage a "simple" RegEx, that is a **Regular Expression**, which represents a sequence of characters, or a *pattern*, that define a search pattern used in several string-searching algorithms or during string-input validation operation; in other word, the aim of a RegEx pattern is to matches a target string.

An example of RegEx is `gr(a|e)y` which matches both `gray` or `grey` strings.

A RegEx pattern is made up of a sequence of elements including:

Atoms which represents a regular character that has a literal meaning.

Meta-characters which represents an element with a very special meaning, depending on context.

The majority of RegEx processor engines supports up to 14 meta-characters including:

`$ & / () [] ^ @ \ | + - *`

According to documentation, the RegEx processor engine of the SimpleRegex class supports only 2 meta-characters:

- `.` which matches any single character.

This meta-characters is known as *wildcard*. For example, `a.b` matches any string that contains an "a", then any other character and then "b", like `axb`.

- `*` which matches any string.

For example, `a.*b` matches any string that contains an "a", and then the character "b" at some later point, like `a andrea graziani b`.

3.1.1 Methods testing analysis.

3.1.1.1 SimpleRegex.

Obliviously, that method represents the constructor of SimpleRegex class and takes up to 2 parameters:

expr A string object containing the RegEx.

caseInsensitive To enable or disable case sensitive.

All valid and invalid equivalence classes are reported in **table 11**, while method signature is shown below, in **listing 9**.

From identified equivalence classes, we have developed $1 \times 2 = 2$ valid test cases and $1 + 1 + 1 + 1 + 1 = 5$ negative test cases. Precisely, we have tested each negative test case when caseInsensitive is both true and false; therefore there is a total of 10 negative test cases, then 12 total tests, 7 of which failed: a very bad result.

Listing 9: Signature of method SimpleRegex

```
public SimpleRegex(String expr, boolean caseInsensitive)
```

3.1.1.2 matches.

This method is used to check the matching between RegEx expression and target string.

All valid and invalid equivalence classes are reported in **table 12**, while method signature is shown in **listing 10**. We have developed 1 valid test case and 1 negative test case, both run twice using different values of caseInsensitive parameter described previously. All tests passed.

Table 11: Equivalence classes and representatives of SimpleRegex method

| Parameter | Equivalence Classes | Representatives |
|-----------------|---|-----------------|
| expr | A not-null String object which: <ul style="list-style-type: none"> Is <i>not</i> blank, that is <code>expr.length() > 0</code> ^[1] It may contain alphanumeric characters It may contain following meta-characters <i>only</i>: ^[2] <ul style="list-style-type: none"> . * It must not contain any <i>escape character</i> ^[3] | a.*b |
| | A not-null String object which: <ul style="list-style-type: none"> <code>expr.length() = 0</code> | "" |
| | A not-null String object which: <ul style="list-style-type: none"> <code>expr.length() > 0</code> It represents an invalid pattern because it contains both unsupported meta-characters and escape characters. | a.*(r t)b\r |
| | A not-null String object which: <ul style="list-style-type: none"> <code>expr.length() > 0</code> It represents an invalid pattern because it contains unsupported meta-characters and no escape characters. | a* |
| | A not-null String object which: <ul style="list-style-type: none"> <code>expr.length() > 0</code> It represents an invalid pattern because it contains escape characters and no unsupported meta-characters | \t |
| caseInsensitive | <i>iEC</i> ₅ null object. | null |
| | <i>vEC</i> ₁ true | true |
| | <i>vEC</i> ₂ false | false |

¹ According to RegEx specification, an empty pattern is *invalid*, because it either returns empty matches or matches with any string depending on RegEx processor engine. In other words, it has an undefined behaviour. In order to match empty strings, `(?![\s\S])` is used.

² Any other meta-character is considered as invalid because, according to SimpleRegex documentation, they are unsupported.

³ An escape character is a character that invokes an alternative interpretation on the following characters in a character sequence. It is a particular case of meta-characters. For example `\u00A9` is an escape character for `©`. SimpleRegex class doesn't support them.

Listing 10: Signature of method matches

```
public boolean matches(String target)
```

3.2 ClassUtil class testing analysis

ClassLoader objects are a part of the *Java Runtime Environment* that dynamically loads Java classes into the *Java Virtual Machine*. When a class is requested, the class loader tries to locate the class and load the class definition into the runtime using a fully qualified class name.

ClassLoader instances use a so-called *delegation approach*: each instance of ClassLoader has an associated parent class loader. When requested to find a class or resource, a ClassLoader instance will delegate the search for the class or resource to its parent class loader before attempting to find the class or resource itself. A special class loader, called *bootstrap*, is used as a parent of all the other ClassLoader instances.

Normally, the Java virtual machine loads classes from the local file system in a platform-dependent manner. For example, on UNIX systems, the virtual machine loads classes from the directory defined by the CLASSPATH environment variable.

REFERENCES

- [1] [n.d.]. BookKeeper concepts and architecture. <https://bookkeeper.apache.org/docs/4.8.2/getting-started/concepts/>. [Online; accessed 4-October-2020].
- [2] [n.d.]. BookKeeper configuration. <https://bookkeeper.apache.org/docs/4.7.3/reference/config/>. [Online; accessed 6-October-2020].

- [3] [n.d.]. The BookKeeper protocol. <https://bookkeeper.apache.org/docs/4.8.2/development/protocol/>. [Online; accessed 5-October-2020].
- [4] [n.d.]. Interface EnsemblePlacementPolicy. <https://bookkeeper.apache.org/docs/4.8.2/api/javadoc/org/apache/bookkeeper/client/EnsemblePlacementPolicy.html>. [Online; accessed 5-October-2020].
- [5] Aalok Ahluwalia, Massimiliano Di Penta, and Davide Falessi. 2020. On the Need of Removing Last Releases of Data When Using or Validating Defect Prediction Models.
- [6] Y. Jia and M. Harman. 2011. An Analysis and Survey of the Development of Mutation Testing. *IEEE Transactions on Software Engineering* 37, 5 (2011), 649–678.
- [7] A.P. Mathur. 2013. *Foundations of Software Testing*. 2/e. Pearson Education India. <https://books.google.it/books?id=OUA8BAAQBAJ>
- [8] Andreas Spillner, Tilo Linz, and Hans Schaefer. 2011. *Software Testing Foundations: A Study Guide for the Certified Tester Exam* (3rd ed.). Rocky Nook.
- [9] M. van Steen and A.S. Tanenbaum. 2017. *Distributed Systems*, 3rd ed. distributed-systems.net.

Table 12: Equivalence classes and representatives of matches method

| Parameter | Equivalence Classes | Representatives |
|-----------|---|------------------|
| target | A not-null String object which <ul style="list-style-type: none"> • <code>expr.length() ≥ 0</code> • Can contain any character. <i>vEC</i> ₁ | "Andrea1234\t= " |
| | <i>iEC</i> ₁ null object. | null |

Table 13: Equivalence classes and representatives of toClass method

| Parameter | Equivalence Classes | Representatives |
|-----------|--|---------------------------------------|
| str | A not-null String object which: <ul style="list-style-type: none"> • Is <i>not</i> blank, that is <code>str.length() > 0</code> • Must be a <i>binary name</i> as defined by <i>The Java™Language Specification</i>.^[1] • Specified <i>binary name</i> must matches an existent and readable .class (or .jar) file. In particular: <ul style="list-style-type: none"> – Into specified str is represent a dollar signs ('\$') used to separate inner and outer classes <i>vEC</i> ₁ | "javax.swing.JSpinner\$DefaultEditor" |
| | A not-null String object which: <ul style="list-style-type: none"> • <code>str.length() > 0</code> • Must be a <i>binary name</i> as defined by <i>The Java™Language Specification</i>. • Specified <i>binary name</i> must matches an existent and readable .class (or .jar) file. In particular: <ul style="list-style-type: none"> – It is a valid <i>binary name</i> containing only periods ('.') used to identify the package name. <i>vEC</i> ₂ | "java.lang.String" |
| | A not-null String object which: <ul style="list-style-type: none"> • <code>str.length() > 0</code> • Must be a <i>binary name</i> as defined by <i>The Java™Language Specification</i>. • Specified <i>binary name</i> must matches an existent and readable .class (or .jar) file. In particular: <ul style="list-style-type: none"> – Specified str represents an array classes with one or more brackets '[' and ']') <i>vEC</i> ₃ | "java.awt.Point[][]" |
| | A not-null String object which: <ul style="list-style-type: none"> • <code>str.length() = 0</code> <i>iEC</i> ₁ | " |
| | A not-null String object which: <ul style="list-style-type: none"> • <code>str.length() > 0</code> • It represents an invalid <i>binary name</i>. <i>iEC</i> ₂ | "\$&andrea..invalid..binary*-"+ |
| | <i>iEC</i> ₃ null object. | null |
| resolve | <i>vEC</i> ₁ true ^[2] | true |
| | <i>vEC</i> ₂ false | false |
| loader | A not-null ClassLoader object which: <ul style="list-style-type: none"> • Must be able to load specified .class file; that is, loader's class paths must include a directory containing specified str class file. <i>vEC</i> ₁ | (see the code) |
| | <i>vEC</i> ₂ null object. ^[3] | null |
| | A not-null ClassLoader object which: <ul style="list-style-type: none"> • Is unable to load specified .class class file. <i>iEC</i> ₁ | (see the code) |

¹ According to the *The Java™Language Specification*, any class or interface must be named by its *binary name*, which must meet several constraints.

Examples of valid class names include:

- java.lang.String
- javax.swing.JSpinner\$DefaultEditor
- java.security.KeyStore\$Builder\$FileBuilder\$1

For details, please visit: <https://docs.oracle.com/javase/specs/jls/se8/html/jls-13.html#jls-13.1>

² When true, the specified class will be *initialized* if it has not been initialized earlier. Initialization of a class consists of executing its static initializers and the initializers for static fields declared in the class.

For details, please visit: <https://docs.oracle.com/javase/specs/jls/se7/html/jls-12.html>

³ Specified ClassLoader object can be null because, in that case, bootstrap class loader would be called.

Table 14: Equivalence classes and representatives of `getClassName(String)` and `getPackageName(String)` methods

| Parameter | Equivalence Classes | Representatives |
|-----------|---|---------------------------------------|
| fullName | A not-null String object which: <ul style="list-style-type: none"> Is <i>not</i> blank, that is <code>str.length() > 0</code> Must be a <i>binary name</i> as defined by <i>The Java™Language Specification</i>. In particular: <ul style="list-style-type: none"> Into specified <code>str</code> is represent a dollar signs ('\$') used to separate inner and outer classes <i>vEC</i> ₁ | "javax.swing.JSpinner\$DefaultEditor" |
| | A not-null String object which: <ul style="list-style-type: none"> <code>str.length() > 0</code> Must be a <i>binary name</i> as defined by <i>The Java™Language Specification</i>. In particular: <ul style="list-style-type: none"> It is a valid <i>binary name</i> containing only periods ('.') used to identify the package name. <i>vEC</i> ₂ | "java.lang.String" |
| | A not-null String object which: <ul style="list-style-type: none"> <code>str.length() > 0</code> Must be a <i>binary name</i> as defined by <i>The Java™Language Specification</i>. In particular: <ul style="list-style-type: none"> Specified <code>str</code> represents an array classes with one or more brackets '[' and ']') <i>vEC</i> ₃ | "java.awt.Point[][]" |
| | A not-null String object which: <ul style="list-style-type: none"> <code>str.length() = 0</code> <i>iEC</i> ₁ | " |
| | A not-null String object which: <ul style="list-style-type: none"> <code>str.length() > 0</code> It represents an invalid <i>binary name</i>. <i>iEC</i> ₂ | "\$%&andrea..invalid..binary*--+" |
| | <i>iEC</i> ₃ null object. | null |

Table 15: Equivalence classes and representatives of `getClassName(Class)` and `getPackageName(Class)` methods

| Parameter | Equivalence Classes | Representatives |
|-----------|--|-----------------|
| cls | <i>vEC</i> ₁ A not-null Class object. | "andrea.Test" |
| | <i>iEC</i> ₁ null object. | null |