

Table of Contents

Development	1
Unit Tests	1
Test Driven Development	5
Pair Programming	7
Code Reviews	8
Static code analysis (backend)	9
Integration tests	12
Easy infrastructure setup	12
Easy application setup	12
Concurrency in application code accounted for	12
GUI Style Guide defined	12
Application Monitoring	13
Scalability requirements known and accounted for	13
Performance requirements known and accounted for	13
Application events logging	13
OWASP Top 10 in Definition of Done	13
Authorization model defined	13
Continuous Integration	13
Continuous Delivery	13
Continuous Deployment	13
Documentation tracked in VCS	13
Documentation generated during CI	13
Parts of the documentation generated automatically	13
Automatic documentation of the executed tests	13
Documentation scope agreed	13
JS application framework	13
JS Build process	13
JS modules dependency management	13
JS Unit test	14
CSS builder	14
Static code analysis (JavaScript)	14
Truly RESTful interfaces	14
HTML validation	14
Database schema versioning	14
Database data versioning	14
Concurrency for DB writes	14
Version Control System	14
Branching strategy	14

Development

Unit Tests

View of Andrzej Bednarz

Introduction

Unit testing is at the core of engineering practices. It's not just a practice, it is a foundation on which many other, more sophisticated practices are built. Without unit tests it would be impossible to apply techniques like [Continuous Delivery](#) or [Test Driven Development](#) and it would make [code refactorings](#) and generally code maintenance much harder and error prone.

What do we understand by unit tests?

There are many definitions of "unit test" in the industry, see e.g. [Martin Fowler - UnitTest](#), which makes the term rather confusing. What we understand by unit test is simply:

- they are written by programmers themselves
- they run fast, in seconds rather than minutes
- they are fine-grained, each test verifies single "thing" (in other words - there is one reason to fail).

Implementation details of how we write unit tests vary, depending on the tools and technologies the project is using. Nevertheless, common goal is to end up with a system that has complete suite of tests which "cover" all functionalities of the system. Thanks to that we are able to clean the code and to improve it anytime, not worrying that we accidentally break something. After each change we can run our tests and make sure that all works as expected. That's really powerful. It leads to higher quality of the code, fewer bugs, faster development, happier programmers and - at the end - more satisfied customers.

It also means that tests are not something additional, optional. They are required and are very important part of the system.

Testability

Writing tests might be a challenge, especially for old, legacy code where no tests were so far developed. Therefore when we start to develop a system, we pay a lot of attention from the very beginning to write testable code. Rules for writing testable code are described in many places, e.g. [Guide to testable code by Misko Hevery](#), but details depend on project technologies, programming language paradigms, etc. As a general rule, we try to think about tests while writing production code or even to write tests before we implement a new feature. Strictly following that practice is called TDD and is described in chapter [Test Driven Development](#). That would automatically ensure testability.

Even in legacy code without tests, we try to write tests firstly, before we make any changes. Thanks to

that we can be sure that we didn't break anything while doing a bug fix or new feature development.

How we are doing unit testing?

The approach we take for writing tests is to treat tests as executable specification. It simply means that tests describe desired functionality. Each test verifies one, single behavior of the system. Since tests focus on behavior, they are not bound to implementation details. They verify public API of a "unit", ignoring private (hidden) methods and internal implementation structure.

Let's see an example. If we need to write a function that returns filtered and sorted list of items, in a test we try to verify desired behavior only by public API, ignoring the fact that filtering and sorting internally is implemented in separate functions.

Unit tests

```
@Test
public void shouldFilterActiveOnly() {
    MyRepo repo = new MyRepo();

    List<Item> items = repo.findItems();

    assertThat(items).are(activeOnly());
}

@Test
public void shouldSortList() {
    MyRepo repo = new MyRepo();

    List<Item> items = repo.findItems();

    assertThat(items).isSorted();
}
```

Example of implementation

```
public class MyRepo {

    public List<Item> findItems() {
        List<Item> items = new ArrayList(new Item(), new Item());
        List<Item> filteredItems = filter(items);
        return sort(filteredItems);
    }

    private List<Item> sort(List<Item> items) {
        List<Item> itemsSorted = new ArrayList(items);
        itemsSorted.sort(null);
        return itemsSorted;
    }

    private List<Item> filter(List<Item> items) {

        List<Item> filtered = new ArrayList<>();

        for (Item item : items) {
            if (item.isActive()) {
                filtered.add(item);
            }
        }
        return filtered;
    }
}
```

As we see, tests use only public method `findItems` to verify both filtering and sorting and are independent of implementation details, they ignore two private functions: `filter` and `sort`. What advantages does it have? Well, firstly, all functionalities are "covered" by the tests. Secondly, we are not bound to implementation details and these details can evolve without breaking tests. For instance, in Java 8 you do filtering and sorting using new Stream API in a single chain of function calls:

Implementation in Java 8

```
public List<Item> findItems() {
    List<Item> items = new ArrayList(new Item(), new Item());
    return items.stream()
        .filter(Item::isActive)
        .sorted()
        .collect(Collectors.toList());
}
```

We made quite a big change, we removed 2 private methods but these are purely implementation

details. No tests require re-implementation. We are free to refactor, to improve technical details, yet the tests remain clean and stable because they are focused on **behavior**.

The tests that we presented show also basic structure of the test, which consists of 3 main steps: *Given*, *When*, *Then*, as on the example below:

Structure of a test

```
@Test
public void shouldSortList() {
    MyRepo repo = new MyRepo(); # <1>

    List<Item> items = repo.findItems(); # <2>

    assertThat(items).isSorted(); # <3>
}
```

- ① Section *Given* - should set up all required objects, pre-conditions to the test
- ② Section *When* - behavior, action you are testing
- ③ Section *Then* - verifies expected state

Ideally, tests have only these 3 lines. Even if there is more to do in a test, it is always possible to refactor to these 3 lines.

We also do, from time to time, other kinds of testing which also employ unit test tools (like JUnit). These could be for instance learning tests (that verify how external library works), low level detailed tests which drive design of implementation details. Nevertheless they are not obligatory, they might be deleted when no longer needed, so not all rules mentioned above apply here. We also sometimes employ unit testing tools for integration tests which is described in a separate article [Integration tests](#).

References

Online resources:

1. <http://martinfowler.com/bliki/UnitTest.html>
2. <http://blog.8thlight.com/uncle-bob/2013/09/23/Test-first.html>
3. <http://blog.8thlight.com/uncle-bob/2014/01/27/TheChickenOrTheRoad.html>
4. <http://blog.arkency.com/2014/09/unit-tests-vs-class-tests/>

Books:

1. <http://www.amazon.com/Clean-Code-Handbook-Software-Craftsmanship/dp/0132350882>
2. <http://www.amazon.com/Test-Driven-Development-Kent-Beck/dp/0321146530>

Test Driven Development

View of Andrzej Bednarz

Introduction

TDD is a way of developing software in which you write tests first, before implementing a functionality. What kind of tests? It's about unit tests as defined in article [Unit Tests](#). However, TDD additionally requires to apply a set of rules (details are below). Thanks to TDD you get source code that is fully testable, fully covered by tests, code that you can trust with your life. It means you know exactly what your code is doing (tests are executable spec), you won't be afraid to clean code, improve it, introduce any change. Your team is able to consistently go fast. Sounds great, doesn't it? It is not trivial to write code in TDD manner though, but let's firstly see what exactly we mean by TDD.

How to do TDD?

Apart from the general rule that you write tests before any production code, there are slight differences in details, of how you do TDD. We like Uncle's Bob approach most. He defines TDD as an activity performed in 4 cycles, see [TheCyclesOfTDD](#).

Firstly, you are not allowed to write a single line of production code without failing test. This is the lowest level rule expressed by so-called [Three Laws of TDD](#):

IMPORTANT

1. You must write a failing test before you write any production code.
2. You must not write more of a test than is sufficient to fail, or fail to compile.
3. You must not write more production code than is sufficient to make the currently failing test pass.

Secondly, since *Getting software to work is only half of the job* (Kent Beck), we need to clean (via refactoring) the code. Therefore once you completed a unit test (or a set of small tests), you need to have a look at code you developed from a little distance and think what to improve having long term maintainability in mind. This way of looking at development cycle is called *Red-Green-Refactor*.

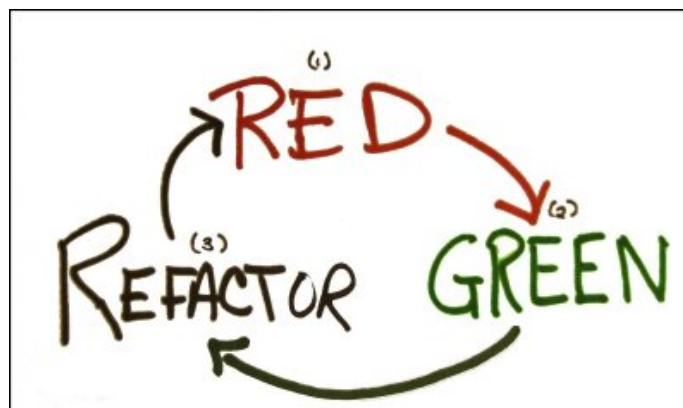


Figure 1. Red-Green-Refactor cycle

Please note that it means we don't have any separate phase in the project called *refactoring* or *code clean up*. Refactorings and care about code quality are continuous, inherent in development process.

Thirdly, even applying previous rules, it's easy to fall into troubles. You may come up with a test that forces you to write tons of production code or even to throw away your whole current implementation. Instead, you should work in an incremental, step-by-step manner. How we can achieve that? By making our implementation code more and more generic with every subsequent test. Our code should not only fulfill requirements imposed by tests, it should naturally flow into direction of generic solution for all possible tests. Also tests themselves could break this natural process. If a test requires *revolution* in code maybe it's worth considering a smaller step. This cycle is called [Specific/Generic cycle](#) and should be considered every couple of tests. It definitely requires some practice and skills to get it right.

Fourthly, while writing small test cases and small portions of code, it might be easy to lose the big picture. Therefore from time to time (say every couple of hours), we need to consider also if our code is in line with general architecture outline, e.g. if we don't call DB directly from a GUI component. Generally, our architectures tend to follow [Clean Architecture](#) principles and as such lead to universal solutions which are highly testable, provide separation of concerns, clearly define boundaries of the system and its internal components.

In TDD tests are equally important as production code, therefore the same amount of care and attention is paid to them.

When to use TDD?

We are not dogmatic about TDD. It is regarded as a good practice and if a team does not follow it, it must have good reason for it. Having said that, our experience is that some areas of software development are very well suited for TDD (algorithmics, finding new solutions), for others it is not very pragmatic (legacy systems, standardized issues - where the implementation is already known) and sometimes even not really possible (front-end stuff like CSS). Nevertheless we **always** write our code having testability in mind and advise to consider TDD in projects.

References

Online resources:

1. <http://blog.cleancoder.com/uncle-bob/2014/12/17/TheCyclesOfTDD.html>
2. <http://blog.8thlight.com/uncle-bob/2013/09/23/Test-first.html>
3. <http://martinfowler.com/bliki/TestDrivenDevelopment.html>
4. http://www.jamesshore.com/Agile-Book/test_driven_development.html

Books:

1. <http://www.amazon.com/gp/product/0321146530>

Pair Programming

View of Andrzej Bednarz

Agile software development requires, among other things, frequent feedback on all possible levels. There is no better way to get early feedback during software development than pair programming, see figure below.



Figure 2. Feedback loops in Agile on XP example

Pair programming is the lowest possible level feedback loop. It provides instant feedback from a developer to a developer while working on code. Sitting in pairs while programming has proved to lead to higher quality code, better knowledge sharing and mutual encouragement (see [Pair Programming in Wikipedia](#)). We do pair programming and we found that technique very useful, especially under following circumstances:

- Development of an unclear/challenging requirement
- Two people with different set of skills
- New person in a team
- New technologies and/or unknown tools

At the same time we found that technique might be inefficient for standard and well understood tasks, when it was obvious what needs to be done or when two people were equal in skills and experience. Also remote work - which is becoming more and more popular these days - makes working in pairs a little harder. Therefore we usually do pair programming for short periods of time, when it makes most sense. Rest of the time we find [Code Reviews](#) sufficient.

Online references:

- <http://powersoftwo.agileinstitute.com/2015/02/benefits-of-pair-programming-revisited.html?m=1>
- <http://blog.jayfields.com/2011/08/life-after-pair-programming.html>
- <http://www.extremeprogramming.org/rules/pair.html>
- <http://guide.agilealliance.org/guide/pairing.html>

Code Reviews

View of Andrzej Bednarz

It is obligatory to do either [Pair Programming](#) or code reviews. Both techniques lead to higher quality, fewer bugs, sharing of knowledge and good practices among team members. Therefore, even though they require some investment of time, they lead in fact to higher productivity and lower total cost of the project, see [Wikipedia](#) for references to empirical studies.

Some time ago the dominating type of code reviews was "over the shoulder". It meant that reviewer had to come to developer's machine as the developer walked through the code. Nowadays the standard way to do code reviews is to use [Git SCM](#) together with some sort of assist tools. Firstly, code developed is usually using [Git feature branches](#), which makes very easy for the reviewer to get the changed source code and check it in an isolated environment. Secondly, many tools emerged which mimic [GitHub pull requests](#). These tools, e.g. [Atlassian Stash](#), make use of Git feature branches and provide convenient web interface for both defining and reviewing pull requests. With these tools you can not only see changes to review in a commit-by-commit view, but also you can share comments, ask developer additional questions or even fire a build on CI environment (e.g. Jenkins) with tests, static code analysis, etc. They also allow to define mini-workflows, e.g. "two positive reviews are required before the change can be merged to the main branch". The following picture shows a basic screen for Atlassian Stash.

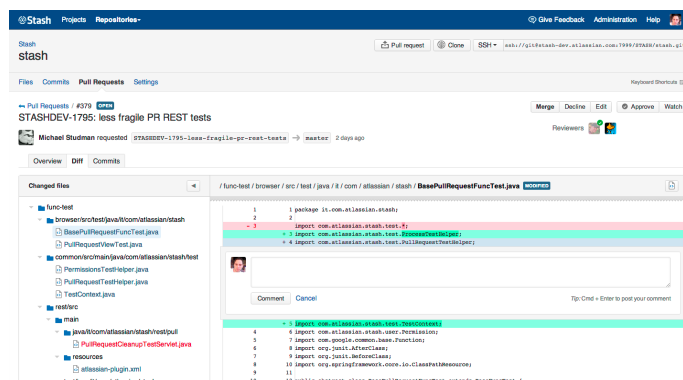


Figure 3. Stash screenshot

Of course tools do not make code review good, they just facilitate it. Teams need to define what exactly need to be checked on code reviews, especially which standards, conventions must be followed (code styles, architecture, GUI guidelines, performance impact, etc.). More on this in subsequent chapters.

Static code analysis (backend)

View of Andrzej Bednarz

Why

There are three main reasons why we use static code analysis in our projects:

- to find potential bugs
- to preserve coding conventions, architecture principles and good practices
- to facilitate reasoning about the code

Characteristic

Implementation details vary across projects and technologies but we found that chosen solution should have the following characteristics:

- Analysis should run automatically as a part of Continuous Integration process.
- It should also be possible to run analysis on developers machines, integrated with IDE preferably.
- Coding style and architecture evolve in time, therefore it should be easy to change and adapt the rules of static code analysis.
- The rules definitions should be maintained centrally - so that all team members share exactly the same settings.
- Results of the analysis should be easy to find and to understand, visual preferably.
- It should be possible to view static code analysis results for deltas, e.g. what issues were introduced in the last commit.
- It should be possible to automatically notify the team if some crucial rules are violated.

If a solution doesn't fulfil most of these requirements, it tends to be ignored by the developers. We also found just limited usefulness of that kind of tools in very small, well organized teams, especially if all members are seniors and proper code review process is in place.

What

When we set up static code analysis, we try to cover the following issues:

- Common conventions for used programming languages
- Common bugs and pitfalls with applied technologies, including security issues

Code duplication

- Design and architecture:
 - Cyclomatic complexity
 - Internal and external dependencies
 - Layering and architecture principles (e.g. GUI layer should not call DB directly.)
- Code coverage . [Strictly speaking, code coverage is a part of dynamic code analysis, nevertheless it is traditionally considered here. We take code coverage as a supportive measure to improve examination of overall code quality. Solely code coverage measure says almost nothing about code quality - i.e. code coverage can be high but code still is not properly tested.]
- Frequency of source code changes per module/class

Example - SonarQube

In the Java world we tried many different solutions, e.g. Checkstyle, FindBugs, PMD, Cobertura, CodePro AnalytiX but recently decided to use [SonarQube](#) as a standard tool. If more sophisticated solution is required, paid tool [JArchitect](#) is a good replacement.

On the screenshots below, you can see how the aforementioned metrics and checks are implemented with SonarQube.

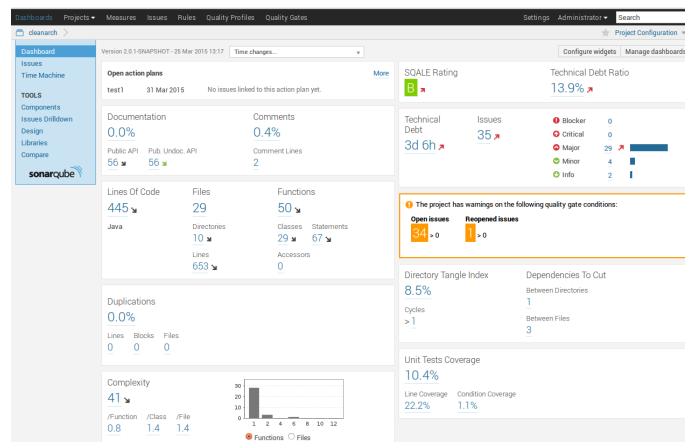


Figure 4. SonarQube dashboard

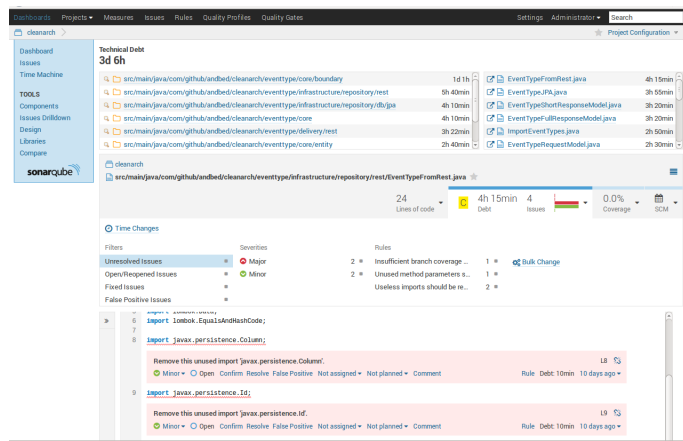


Figure 5. Issues caught by SonarQube

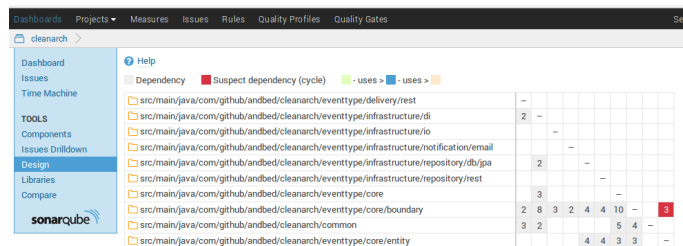


Figure 6. Dependencies of modules and a cycle indicated in red, SonarQube

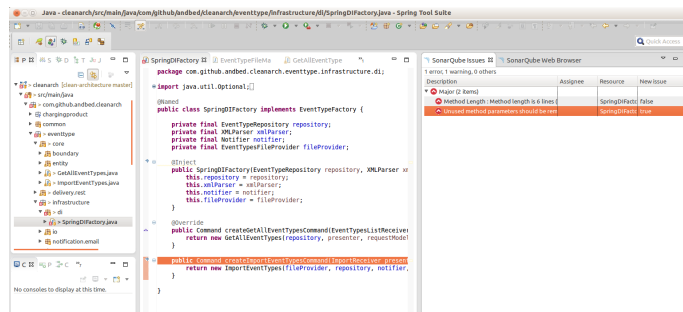


Figure 7. SonarQube integration with Eclipse

Additionally, we find the following features of SonarQube helpful:

- Predefined reasonable set of rules, including FindBugs, Checkstyle, PMD
- Smooth integration with Jenkins - a CI server
- Smooth integration with Eclipse and IntelliJ IDEA.
- Centralized and easy rules management
- Visual tools for analysis: dashbord, drill-down, time machine, incremental (deltas) analysis
- Alerting
- Open source - can be extended and adapted (e.g. rules).

Unfortunately, SonarQube (as probably every tool) does not provide full range of metrics we find valuable. We deal with such situations using some other ways, often simple code investigations. For instance SonarQube says nothing about frequency of code changes. It is an important indicator showing "suspicious" places in code which could be badly designed or could contain a lot of bugs. In such case when working with GIT we use the following simple method:

History of source code changes

```
# To get a list of all resources sorted by number of changes:

find . -path ./git -prune -o -name "*" -exec sh -c 'git log --follow --format=oneline $1
| wc -l | awk "{ print \$1, \"\\t\\t\\t\\t\", \"$1\\t\" }" ' {} {} \; | sort -nr

# To get a list of all changes for give resource (e.g. a directory)

git log resource_name

# Thanks to these simple commands we could draw conclusions which places of the system
are changed most frequently and what the histogram is. A lot of changes at the beginning
is fine, but equally distributed throughout the lifetime might mean something bad.
```

References

- <http://www.sonarqube.org/>, SonarQube home page
- <https://www.youtube.com/watch?v=KaLROwp-VDY>, Greg Young - "How to get productive in a project in 24h"

Integration tests

Easy infrastructure setup

from nothing to running in <1h

Easy application setup

from nothing to running in <1h

Concurrency in application code accounted for

GUI Style Guide defined

Application Monitoring

Scalability requirements known and accounted for

Performance requirements known and accounted for

Application events logging

OWASP Top 10 in Definition of Done

Authorization model defined

Continuous Integration

Continuous Delivery

Continuous Deployment

Documentation tracked in VCS

Documentation generated during CI

Parts of the documentation generated automatically

Automatic documentation of the executed tests

Documentation scope agreed

JS application framework

JS Build process

JS modules dependency management

JS Unit test

CSS builder

Static code analysis (JavaScript)

Truly RESTful interfaces

HTML validation

Database schema versioning

Database data versioning

Concurrency for DB writes

Version Control System

Branching strategy