

Identification of Module Boundaries in a Modular Monolith Architecture using Automated technologies

UNIVERSITY OF TURKU
Department of Computing
Master of Science (Tech) Thesis
February 2024
Florian Dejonckheere

UNIVERSITY OF TURKU
Department of Computing

FLORIAN DEJONCKHEERE: Identification of Module Boundaries in a Modular Monolith Architecture
using Automated technologies

Master of Science (Tech) Thesis, 4 p., 4 app. p.
Department of Computing
February 2024

The modular monolith architecture emerged in recent years as the harmonization of the monolithic and microservices architectures. The paradigm offers a compromise between modularity, flexibility, and scalability. Many monolithic applications are being migrated to modular monoliths or microservices entirely, to satisfy increasingly complex and volatile business requirements. This process is labour-intensive, slow, and may take months to years for larger codebases. Modularization of a codebase typically requires the developer to have an intimate knowledge of both the application code and domain.

In this thesis, we investigate the modular monolith software architecture, and how modules are typically determined as part of the modularization efforts. We propose an automated solution based on dependency analysis and machine learning algorithms to aid in the identification of module boundaries, and evaluate its effectiveness using a case study. We discuss the results and draw conclusions about the proposed solution.

Keywords: software architecture, monolith, microservices, modular monolith

Contents

1. Introduction	1
1.1. Scope and goal	1
1.2. Motivation	1
1.3. Methodology	1
1.4. Outline	2
2. Background	3
2.1. Monolith architecture	3
2.2. Microservice architecture	3
3. Related work	4
4. Modular monolith architecture	5
4.1. Background	5
4.2. Challenges and opportunities	5
4.3. Modularization	5
5. Modularization	6
6. Case study	7
6.1. Background	7
6.2. Analysis	7
6.3. Evaluation and results	7
6.4. Discussion	7
7. Conclusion	8
7.1. Future work	8
References	10

List of Figures

List of Tables

List of Acronyms

API	Application Programming Interface
UI	User Interface

1. Introduction

1.1. Scope and goal

This research is centered around three research questions:

Research Question 1: Which challenges and opportunities arise when considering adoption of the modular monolith architecture in an existing codebase?

Research Question 2: Which automated technologies exist to aid modularization of monolithic codebases?

Research Question 3: How can (*automated technology*) effectively identify optimal module boundaries in a modular monolith architecture?

To answer the first research question, we will first define the modular monolith architecture, and examine what sets it apart from monolithic and microservices architectures. Then, we will proceed to investigate the merits and drawbacks of the software architecture when applied to an existing codebase.

For the second research question, we will enumerate the existing technologies to aid modularization of monolithic codebases, and choose one automated technology for further examination. (*Automated technology*) will then be implemented for a given use case, and compared to manual modularization efforts in terms of accuracy, efficiency, development velocity. This comparison will help us to answer the third research question.

The goal of this research can be summarized as follows:

1. Investigate the merits and drawbacks of the modular monolith architecture in an existing codebase
2. Investigate the use of automated technologies to modularize a monolithic architecture

The proposed solution will add value to the field of software engineering, and will be able to be used as a base for future improvements regarding automated modularization of monolith codebases.

1.2. Motivation

1.3. Methodology

To answer the first and second research question, a literature study is conducted. The literature study lists relevant literature, and reviews the history and evolution of modular monolith architectures, and summarizes the state of the art in automated modularization technologies.

Subsequently, the most appropriate automated technology is chosen, and a proposed solution is implemented for a given use case. Based on certain key criteria, the proposed solution is compared to manual modularization efforts.

Finally, the findings are summarized, and an outlook on future work is given.

1.4. Outline

The thesis is divided into three parts.

The first part comprises the background and related work. In Chapter 1, the scope and goal of the research is defined, and the research questions are formulated. The stakeholders are identified, and the methodology is explained. Chapter 2 introduces the reader to the research background and necessary concepts. In Chapter 3, the existing literature is reviewed, and the state of the art is presented.

The second part of the thesis, starting with Chapter 4, is dedicated to the first research question. The modular monolith architecture is defined, and its merits and drawbacks are discussed.

The third part aims to solve the second and third research question. Chapter 5 gives an introduction into the automated modularization of monolith codebases, listing the existing technologies. It then continues to focus on one automated technology, (*automated technology*), and explains its implementation. Chapter 6 applies (*automated technology*) on a given case study, and compares it to manual modularization efforts.

Finally, Chapter 7 summarizes the findings, and gives an outlook on future work.

2. Background

2.1. Monolith architecture

2.2. Microservice architecture

3. Related work

4. Modular monolith architecture

4.1. Background

4.2. Challenges and opportunities

4.3. Modularization

5. Modularization

6. Case study

6.1. Background

6.2. Analysis

6.3. Evaluation and results

6.4. Discussion

7. Conclusion

7.1. Future work

References