

Technische Hochschule Ingolstadt

Specialist area Computer Science

Bachelor's course Computer Science

Bachelor's thesis

Subject: Conception, implementation and evaluation of a highly scalable and highly available Kubernetes-based SaaS platform on Kubernetes Control Plane (KCP)

Name and Surname: David Linhardt

Issued on: TODO: Insert Issue Date

Submitted on: TODO: Insert Submit Date

First examiner: Prof. Dr. Bernd Hafenrichter

Second examiner: TODO: Insert Second examiner

Abstract

Contents

1	Introduction	3
1.1	Problem Statement and Motivation	3
1.2	Objectives and Scope	3
1.3	Structure of the Thesis	3
2	Fundamentals	3
2.1	Kubernetes and Multi-Tenancy	3
2.2	Kubernetes Control Plane (KCP)	3
2.3	SaaS Architecture and Automation	3
3	State of the Art and Related Work	3
3.1	Zero-Downtime Deployment Strategies	3
3.2	Kubernetes Scaling Methods	3
3.3	Multi-Tenancy Concepts in the Cloud	3
4	Conceptual Design	3
4.1	System Requirements	3
4.2	Architecture Design with KCP for SaaS	3
4.3	Automated Deployment Strategies	3
5	Prototypical Implementation	3
5.1	Infrastructure with KCP	3
5.2	Tenant Provisioning	3
5.3	Scaling Mechanisms	3
5.4	Monitoring and Logging	3
6	Evaluation	3
6.1	Performance Measurements	3
6.2	Scaling Scenarios & Optimizations	3
6.3	Discussion of Results	3
6.4	Related Work	3
7	Conclusion and Outlook	3
7.1	Summary	3

Contents

7.2	Personal Conclusion	3
7.3	Future Outlook	3
	References	3
	List of Figures	3

Contents

Glossary

1 Introduction

1.1 Problem Statement and Motivation

1.2 Objectives and Scope

1.3 Structure of the Thesis

2 Fundamentals

2.1 Kubernetes and Multi-Tenancy

2.2 Kubernetes Control Plane (KCP)

2.3 SaaS Architecture and Automation

3 State of the Art and Related Work

3.1 Zero-Downtime Deployment Strategies

3.2 Kubernetes Scaling Methods

3.3 Multi-Tenancy Concepts in the Cloud

4 Conceptual Design

4.1 System Requirements

4.2 Architecture Design with KCP for SaaS

4.3 Automated Deployment Strategies

5 Prototypical Implementation

5.1 Infrastructure with KCP

5.2 Tenant Provisioning (Automation, Multi-Tenancy)

5.3 Scaling Mechanisms (Horizontal Pod Autoscaler)

5.4 Monitoring and Logging (Prometheus, Grafana)

6 Evaluation

6.1 Performance Measurements (Downtime, Latency, Scaling)

6.2 Scaling Scenarios & Optimizations

6.3 Discussion of Results

6.4 Related Work

Appendix