



Cloud Oriented Micro Services Resource Optimization by Content Delivery Networks

P.P.A.S. Jayasundara

(Reg. No.: MS20909482)

M.Sc. in IT

Specialized in Enterprise Application Development

Supervisor: Ms Shashika Lokuliyana

October 2021

**Department of Software Engineering
Faculty of Graduate Studies and Research
Sri Lanka Institute of Information Technology**

Table of Contents

Table of Contents.....	2
List of Figures	3
List of Tables	4
Abstract.....	5
Acknowledgment	6
Chapter 1 Introduction	7
1.1 Background	7
1.1.1 Load Balancer (LB).....	8
1.1.2 Distributed Denial of Service (DDoS)	10
Chapter 2 Objectives.....	15
2.1 Main Objectives	16
Chapter 3 Micro Service.....	17
3.1 Usage of Microservices	18
3.2 Message Communication.....	19
3.3 Message Communication Pattern.....	20
3.4 Communication Guarantees	20
3.5 Fault Tolerance	22
3.6 Service orchestration	23
3.6.1 Leader Election.....	23
3.7 Message Serialization.....	25
Chapter 4 Resource Utilization	28
4.1 Redis Cache	28
4.2 Reactive Programing	29
4.2.1 Non-Blocking Requests	30
4.2.2 Reactive Stream Specifications	30
4.3 Reverse Proxy.....	31
4.4 Message Broker	33
4.5 Message brokers vs. APIs.....	35
Chapter 5 Infrastructure design.....	43
5.1 Kubernetes.....	43
Conclusion.....	47
References	48
Appendix	50

List of Figures

Figure 1 Fundamental Micro-Service Architecture	7
Figure 2 Cloud Load Balance Architecture	9
Figure 3 DDoS Attack Simulator	12
Figure 4 AWS Load Balancer Functionalities	13
Figure 5 Available Technology Stack for Microservice [15]	17
Figure 6 Monolithic Architecture	18
Figure 7 TCP Socket Communication	19
Figure 8 Message Transfer Mechanism	21
Figure 9 Message Checkpoint [17].....	22
Figure 10 Cluster Message Communication	23
Figure 11 Leader Implementation	24
Figure 12 Message Serialization performance (time).....	25
Figure 13 Message Serialization Performance (payload size).....	26
Figure 14 Request Performance with ProtoBuff.....	27
Figure 15 Machine to Machine Serialization Performance.....	27
Figure 16 Redis Cache Throughput	29
Figure 17 Message Broker is Redis Cache.....	37
Figure 18 API Request Endpoint	38
Figure 19 Publish to Redis Stream	39
Figure 20 Redis Stream Consumer Endpoint	40
Figure 21 Redis Cache Server Configuration.....	41
Figure 22 Docker Containerized Image building process.....	41
Figure 23 docker Composer	42
Figure 24 K8 Architecture	45
Figure 25 K8 Deployment Config	46
Figure 26 Project Budget.....	50
Figure 27 Project Plan	50
Figure 28 Configure Node Groups	51
Figure 29 Create Elastic Kubernetes Cluster	52

List of Tables

Table 1 Load Balancer Capacity Unit Calculation..... 14

Abstract

In the field of modern information technology, the most intriguing topic is cloud-based application development. After several decades of rapid development and research on cloud technologies, nowadays almost all cloud service providers are providing a massive range of services with higher reliability but there are a couple of business domains that having additional technical requirements and these are unique to their business domains. Capital Market and Finance is one of such specific business domains which need to address additional technical and compliance requirements.

The main technical barrier in this domain is providing business functionalities for all users across the globe with micro-second level latency. Therefore, when developing and maintaining such a system, we are highly concern about system throughput and hardware resource allocation. While on the subject, the cloud-based system architecture is an ideal infrastructure for this kind of application development because we can upgrade hardware resources within a couple of minutes. however, there are significant issues remains as it is.

- Message queues are growing unexpectedly until resource upgrade.
- Lack of accurate cloud services to identify duplicate API requests.
- User connectivity and API access are limited due to service back off
- Peek time is limited to sort period and resources are billing on hours.
- System recovery in machine terminate is very costly mechanism

As a matter of above technical concerns, we are conducting this research to propose a better solution to handle these types of technical barriers in without upgrading hardware resources unnecessarily and the proposed solution will not be limited to Capital market but it can be used for any application service to utilize their hardware resources while high network traffic

Acknowledgment

A special acknowledgment should be made to my supervisor Ms. Shashika Lokuliyana who has made a great effort in fine-tuning this research analysis and sharing her knowledge and experience to make this research realistic and it is a great privilege to be under her supervision.

Mr. Nuwan Abesiriwardana (Ustocktrade LLC | CTO) who is given the initial idea and the requirement related to this research and valuable ideas for implementation of this project.

This research project and the entire 2 years of MSc program could not have been completed without having invaluable support from various groups specially in this pandemic situation. I would like to thanks to my parents and family members, office colleges and the friends who motivate me sharing their experience when developing large scale cloud base application.

Thank you all for providing me a lot of support and motivate me during my research studies.