Design Document: Architecture and Cluster Sizing

# Table of Contents

#### 1. Introduction

#### 2. Architecture Overview

#### 1. Components

#### 2. Interactions

#### 3. Cluster Sizing

#### 1. Methodology

#### 2. Requirements

#### 3. Sizing Considerations

#### 4. Deployment Strategy

#### 5. Monitoring and Maintenance

#### 1. Monitoring Tools

#### 2. Maintenance Processes

#### 6. Security Measures

#### 7. Conclusion

# 1. Introduction

This document outlines the architecture and cluster sizing considerations for the deployment of a Streaming platform utilizing AWS services in the Stockholm Region. It provides a detailed overview of the components, their interactions, and the methodology used for cluster sizing.

## 2. Architecture Overview

The architecture utilizes AWS services, specifically Amazon EKS for Kubernetes deployment, in the Stockholm Region with three Availability Zones (AZs).

## 2.1 Components

**EKS Cluster**: Utilized for Kubernetes deployment.

**Node Pools**: Three node pools, one for each AZ, consisting of m5.xlarge instances with the following specifications:

- Min: 3 nodes

- Desired: 3 nodes

- Max: 4 nodes

**Brokers**: Two brokers per AZ, distributed across different nodes.

**Shared Services**:

Zookeeper

Schema Registry

Ksqldb

Connect Control Centre

Rest Proxy

## 2.2 Interactions

The brokers are distributed across different nodes within each AZ. Shared services such as Zookeeper, Schema Registry, Ksqldb, and Connect Control Centre are deployed on separate nodes within each AZ.

## 3. Cluster Sizing

Cluster sizing is based on the workload requirements and resource specifications for each service.

### 3.1 Methodology

The sizing methodology considers the resource requests and limits specified for each service, along with current workload in Stockholm region.

### 3.2 Requirements

The resource requests for each service are as follows:

Kafka: 16GB Memory Request, 16GB Memory Limit, 4 mCPU Request

Connect: 6GB Memory Request, 16GB Memory Limit, 2 mCPU Request

Zookeeper: 4GB Memory Request, 8GB Memory Limit, 1 mCPU Request

Control Center: 4GB Memory Request, 8GB Memory Limit, 1 mCPU Request

Schema Registry: 1GB Memory Request, 4GB Memory Limit, 0.5 mCPU Request

Rest Proxy: 1GB Memory Request, 4GB Memory Limit, 0.5 mCPU Request

KsqlDB: 1GB Memory Request, 2GB Memory Limit, 4 mCPU Request

### 3.3 Sizing Considerations

Factors influencing cluster sizing include peak vs. average load, growth projections, redundancy, fault tolerance, and cost considerations.

## 4. Deployment Strategy

The deployment strategy includes provisioning and scaling resources across development, staging, and production environments using AWS services and Kubernetes.

## 5. Monitoring and Maintenance

### 5.1 Monitoring Tools

Control Center is set up in managed-only mode, with all data and metrics sent to DataDog for monitoring purposes. DataDog will be utilized as the primary tool for viewing metrics and other data related to the system's performance.

### 5.2 Maintenance Processes

Monitoring tools and maintenance processes are in place to ensure system health, performance metrics tracking, backups, and updates management.

## 6. Security Measures

Security measures include authentication, authorization, encryption, and compliance with relevant regulations to protect the system from potential threats.

## 7. Conclusion

In conclusion, this design document provides a comprehensive overview of the architecture and cluster sizing considerations for the deployment of the system utilizing AWS services in the Stockholm Region. It outlines the methodology, requirements, and considerations for effective deployment, monitoring, and maintenance of the system.