

Synchronizing Data in a Docker Swarm Cluster with MQTTS

Michele Montesi

E-Mail: michele.montesi3@studio.unibo.it

June 18, 2023

1 Introduction

This document presents the project of synchronizing data in a Docker Swarm cluster using the MQTTS protocol. It will illustrate the application context, the specific problem to be solved, and the implemented solution, including the technologies and configurations adopted.

2 Application Context

The project starts from a Docker Swarm environment, which is a container orchestration system that manages a cluster of Docker nodes. It is also assumed that an MQTT (Message Queuing Telemetry Transport) messaging service has been configured on all nodes of the cluster, known as MQTTS (MQTT over TLS/SSL). The use of MQTTS provides secure and encrypted communication between the nodes of the Swarm cluster.

The goal of the project is to achieve a Docker Swarm cluster where the MQTTS service is running on all nodes of the cluster. Furthermore, it is necessary to ensure data persistence and synchronization between nodes using the `lsyncd` utility.

3 Specific Problem

The specific problem is the need to synchronize data between nodes in the Docker Swarm cluster. Due to the distributed nature of the cluster, with its independent nodes, situations may arise where data is inconsistent or out of sync across the various nodes. It is essential to solve this problem to ensure data integrity and consistency in the Docker Swarm environment.

4 Implemented Solution

To address the problem of data synchronization between nodes in the Docker Swarm cluster, the following solution has been adopted:

4.1 MQTTS

The MQTT messaging service has been configured on all nodes of the Swarm cluster using the MQTTS protocol, which provides secure and encrypted connections. The use of MQTTS ensures confidentiality and authentication during communication between the nodes.

4.2 lsyncd

To ensure data synchronization between nodes in the cluster, the lsyncd (Live Syncing Daemon) utility has been utilized. Lsyncd is a tool that monitors file changes in a directory and automatically propagates them to one or more remote destinations.

A specific synchronization rule has been configured in lsyncd to monitor the data directories within each Docker Swarm node and automatically propagate any changes to the remaining nodes. This way, any modifications made on one node will be automatically and transparently synchronized with all other nodes in the cluster.

4.3 Data Persistence

To ensure data persistence in the Docker Swarm cluster, a Docker volume has been configured for the data directories. Docker volumes allow data to be separated from the container, enabling persistence even in the event of container or node restarts. This ensures that data remains available even in failure or maintenance scenarios.

5 Conclusion

This report presented the project of synchronizing data in a Docker Swarm cluster with MQTTS. The project was motivated by the problem of synchronizing data between independent nodes in the Swarm cluster. The implemented solution, which utilizes the MQTTS protocol for secure communication and the lsyncd utility for automatic data synchronization between nodes, was described. The importance of data synchronization in a distributed environment like Docker Swarm was emphasized to ensure data integrity and consistency across the cluster's nodes.