**DATA TRANSFORMER DOCUMENTATION**

**Technical Documentation**

**<DPDHL-Transformer>**

**Company Name**

**S3 Innovate**

**Date**

**27/07/2023**

Table of Contents

1. [Introduction: 3](#_Toc142414775)
2. [Solution Overview: 3](#_Toc142414776)
3. [System Architecture: 3](#_Toc142414777)
4. [Security & privacy: 4](#_Toc142414778)
5. [Performance and Scalability: 4](#_Toc142414779)
6. [Reliability: 5](#_Toc142414780)
7. [Availability And Maintainability: 5](#_Toc142414781)
8. [Change log: 5](#_Toc142414782)
9. [Release notes: 5](#_Toc142414783)

# Introduction:

The following documentation is based on DGF And EXP technical architecture that provides an overview of the Data Transformer application for file processing. It aims to guide developers and clients in understanding the application design for DGF and EXP shipment.

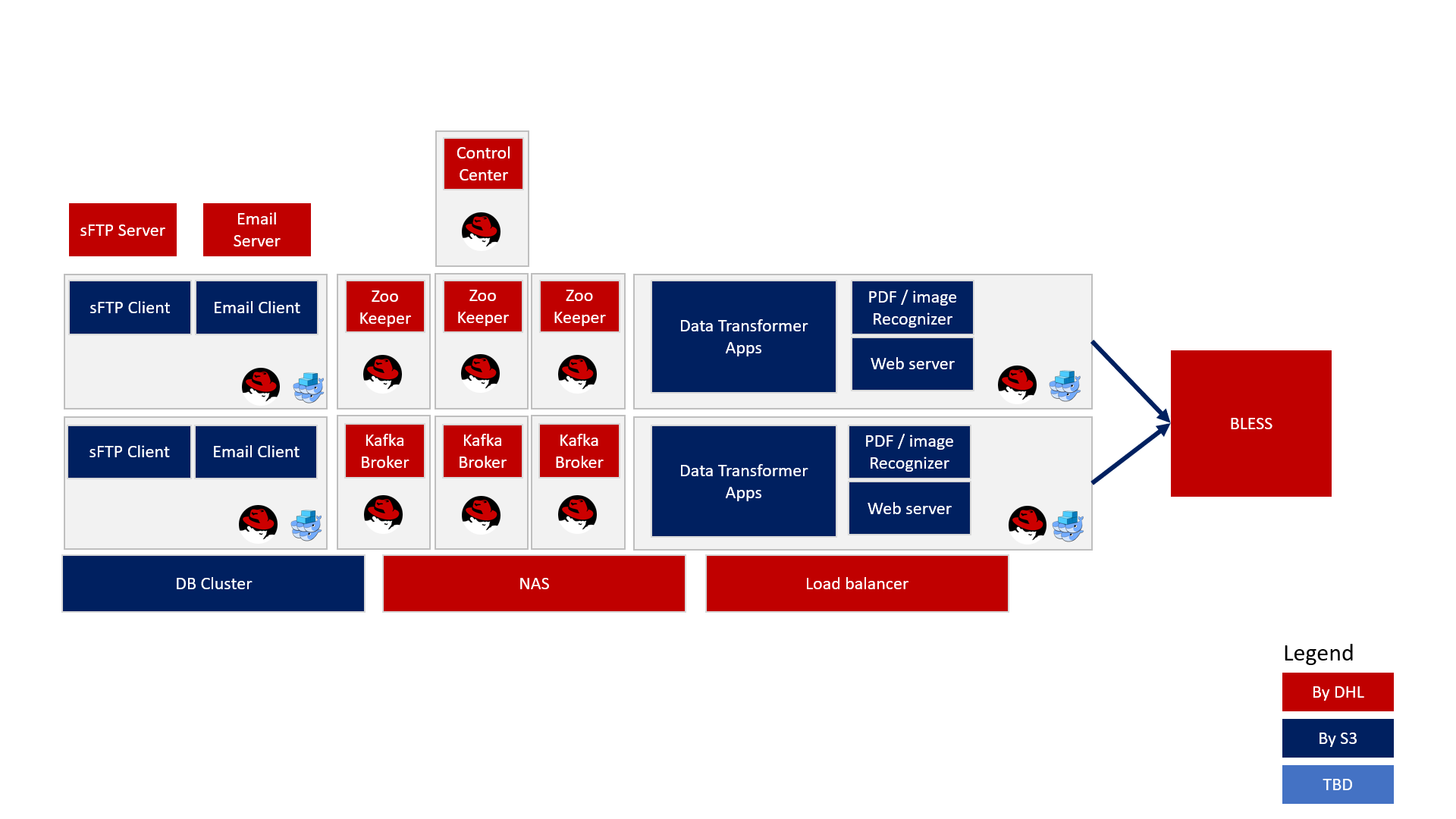
# **Solution Overview**:

The Data transformer application read DGF AND EXP shipment files from the sftp server through Kafka message broker whenever data or file is placed in sftp server. After receive those data mainly it processes the data by using data transformer process mechanism then the application trigger to the bless endpoint with processed data as payload.

# System Architecture:

This diagram will give the overviews of DGF and EXP shipment file processing.

## 



The below key points from the diagram are illustrated below:

* **SFTP Server**: SFTP server mainly implements the SFTP protocol and it allows clients to.

securely upload, download, and manage files on the server. Data transformer systems are responsible for receiving these files from the sftp server through Kafka message broker.

* **Email Client**: The main responsible of email client server is to read the file form the

particular email sender with specific time Interval. And here also Kafka message broker becomes active when it gets file from the email client server. In the diagram we also notice sftp and email client server’s file is being read by Kafka message broker.

* **Kafka-broker**: A Kafka Broker is used for distributed messaging system throughout the

application. In the data transformer application, the Kafka consume the real time data over the different server like sftp or email client server.

# Security & privacy:

To enhance security and restrict access to unprivileged users, the data transformer application employs JWT token authentication. This token is utilized for every request, allowing the application to identify and verify the authenticity of the user.

# Performance and Scalability:

Since the data transformer application for the DGF and EXP shipment file is responsible for processing and transforming data from one format to another, often handling large volumes of data. Here are some considerations and strategies for ensuring good performance and scalability:

Manageable microservices that focus on specific tasks like kafka client microservices for consuming data from sftp server as well as data transformer microservices process those data from kafka. So, it provides independent scaling of different components based on their resource requirements.

The data transformer application utilizes asynchronous processing methos for time-consuming tasks. This allows the microservice to quickly acknowledge requests and then process under the thread for freeing up resources to handle more requests.

The data transformer application uses efficient algorithms and data structures for data transformation to reduce processing time.

# Reliability:

The data transformer application For DGF and EXP utilize containerization like as Docker to isolate microservices and limit the impact of failures to individual components.

This application set up automated recovery processes that can detect failures and take corrective actions. This could involve restarting failed instances, provisioning new instances, or rolling back to a previous version of the microservice.

The data transformer application implements comprehensive error handling mechanisms and logging to capture and analyze errors, making it easier to diagnose and resolve issues.

Achieving reliability for data transformer application is an ongoing process that requires continuous monitoring, testing, and improvement.

# Availability And Maintainability:

For DGF and EXP,The Data transformer application is integrated with number of components that can impact deployment time. After developing the new features or the modifying existing module for this application it takes a little bit time to deploy in the server since it removes the existing image and then it pulls from the docker hub for the changes.

The data transformer application also maintains version control like v1 to v2 for the application code, configurations, and infrastructure scripts. This enables easy tracking of changes and rollbacks.

The application is different from v1 to v2. So, it keeps the code isolated from one version to another.

# 

# Change log:

Since the data transformer application has been upgraded from v1 to v2,the new features has been added in v2.

Minimal api endpoint like /api/v2/transformfile has been introduced to process the file locally to test the file processing. In previous version 1 we have dependency on Kafka as well as sftp server to test the file processing.

To Simplify configuration for Kafka, mapper as well as mapper template, the dedicated configuration path has been isolated from one directory to another.

In the previous v1 the application produced token in data transformer service, Now the Kafka client service is responsible for provide token after file processing completely.

# Release notes:

* Version 2.0 - Released on around 2022.
* The new features or module like Kafka.IBM,Kafka.DutyDrawback,Kafka.Kenya, Kafka.Rwanda has been introduced in v2.
* To enhance clients flexibility for IBM invoice , the Kafka IBM service has been introduced to read the file from the email and sent to the mail with log after processing file.