

Implementing a Metadata-Driven Data Lakehouse on Azure with CF.Cumulus

Leveraging ready-to-go pipelines, connections and metadata to accelerate the delivery process of a cloud data platform.

<https://cloudformations.org>

contactus@cloudformations.org

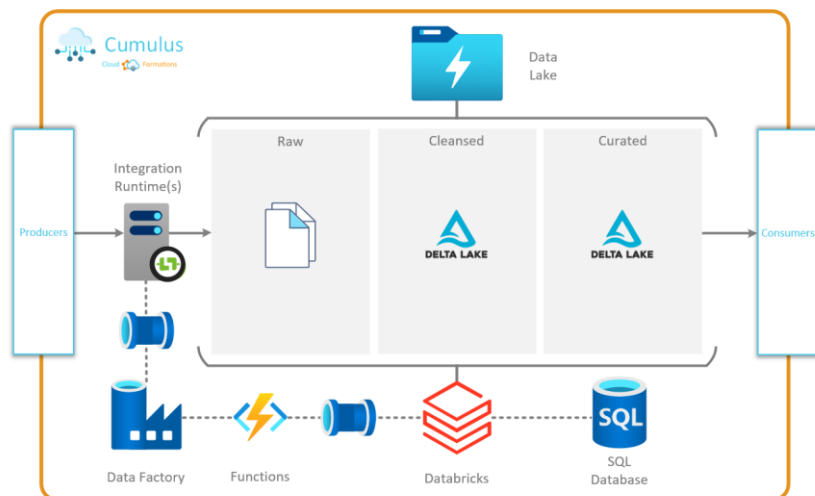
[In/CloudFormations](https://www.linkedin.com/company/cloudformations)

[@CloudFormsLtd](https://twitter.com/CloudFormsLtd)

[CloudFormationsLtd](https://www.facebook.com/CloudFormationsLtd)



Cloud Formations utilised our open-source metadata-driven Lakehouse data platform accelerator, CF.Cumulus, to rapidly ingest data from Oracle and Dynamics 365, leveraging Microsoft Analytics Platform technologies to deliver use cases for a large London based housing association.



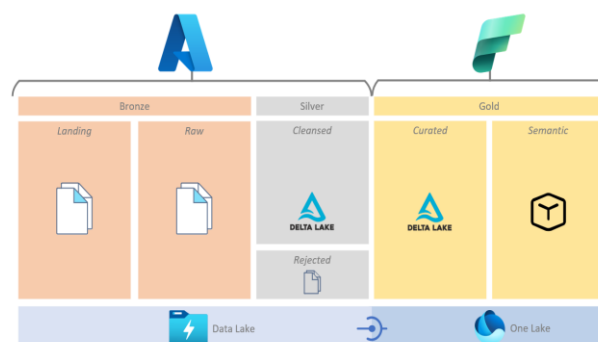
Technologies

Using Microsoft Azure cloud native resources, a data mesh architecture is being designed using the following compute and storage components:

- Data Lake
- Data Factory
- Databricks
- SQL Database
- Functions
- Microsoft Fabric
- Azure DevOps

Solution Overview

Using the CF.Cumulus accelerator provided rapid deployment of a best practice Lakehouse data platform with seamless integration, meeting networking and governance requirements of the client. The pre-built product connections and out-of-the-box CF.Cumulus pipelines in Azure Data Factory saved weeks of development effort facilitating deeper collaboration with the client and enabling self-service of the platform.



Data is then served through Microsoft Fabric for a unified business user experience meeting the strategic ambitions of the organisation and enable consumers to explore and self-serve from trusted, curated datasets. The solution uses the best of all products to deliver a complete 'medallion architecture'.

Key Features

- Metadata-driven data ingestion framework.
- Intelligent data loading through CDC logic and data load history.
- Cost-efficient compute implementation for both development and production experiences.
- Resilient and restartable orchestration pipelines.

Outcome

- The solution provided the business with:
- Datasets ingested from source with little configuration efforts by development team.
 - Repeatable framework and training provided to ingest additional data sources.

Next Steps

Reach out to learn more about how we could implement this architecture in your organisation via our contact details (top right) or please book some time in our calendar via the following link: bit.ly/cf-chat

Our Services

- Data Platform Design
- CF.Cumulus Deployment
- Technical Oversight
- Architecture Review
- End to End Training

