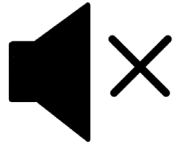


July 18, 2023

CDC Enabled Mass Ingestion Application with Data Lake

- Chandrasekar K, Consulting Manager, IPS

Housekeeping Tips



- Today's Webinar is scheduled for **1 hour**
- The session will include a webcast and then your questions will be answered live at the end of the presentation
- All dial-in participants will be muted to enable the speakers to present without interruption
- Questions can be submitted to "All Panelists" via the **Q&A option** and we will respond at the end of the presentation
- The webinar is **being recorded** and will be available on our **INFASupport YouTube channel** and **Success Portal** - where you can download the **slide deck** for the presentation. The link to the recording will be emailed as well.
- Please take time to complete the **post-webinar survey** and provide your feedback and suggestions for upcoming topics.

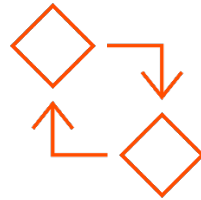
Feature Rich Success Portal



Bootstrap trial and
POC Customers



Enriched Customer
Onboarding
experience



Product Learning
Paths and Weekly
Expert Sessions

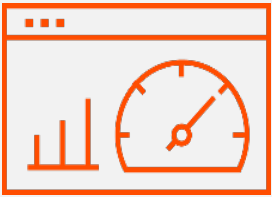


Informatica
Concierge



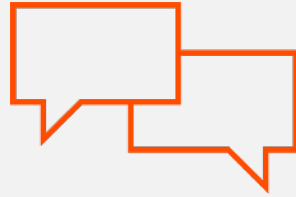
Tailored training and
content
recommendations

More Information



Success Portal

<https://success.informatica.com>



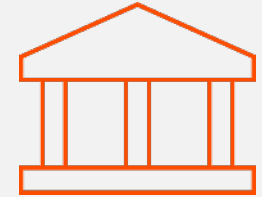
Communities & Support

<https://network.informatica.com>



Documentation

<https://docs.informatica.com>



University

<https://www.informatica.com/in/services-and-training/informatica-university.html>

Safe Harbor

The information being provided today is for informational purposes only. The development, release, and timing of any Informatica product or functionality described today remain at the sole discretion of Informatica and should not be relied upon in making a purchasing decision.

Statements made today are based on currently available information, which is subject to change. Such statements should not be relied upon as a representation, warranty or commitment to deliver specific products or functionality in the future.

18-07-2023



Mass Ingestion Application with Datalake

Chandrasekar.K , Consulting Manager
chak@informatica.com



Agenda

1

Introduction

2

Scope and
Purpose

3

Data Lake

4

Application
Ingestion Task

5

Demo

6

Q&A

Introduction

- This webinar is intended for all IICS developers, administrators, and architects.
- The session will help you learn how to ingest data into datalake using CMI and understanding Application Mass ingestion Usecase.

Scope

- Data load into Datalake (S3)
- Application Ingestion service is used to load data into datalake.
- Leverage Initial and Incremental loads.
- CDC Options
- Recovery from the failure.

Purpose

What is the purpose of Application Mass Ingestion?

- No Code or Wizard Based Template
- CDC enable.
- Realtime Data movement
- Efficiently Ingest & Synchronize your data with scale and automation
- Monitoring Dashboard.

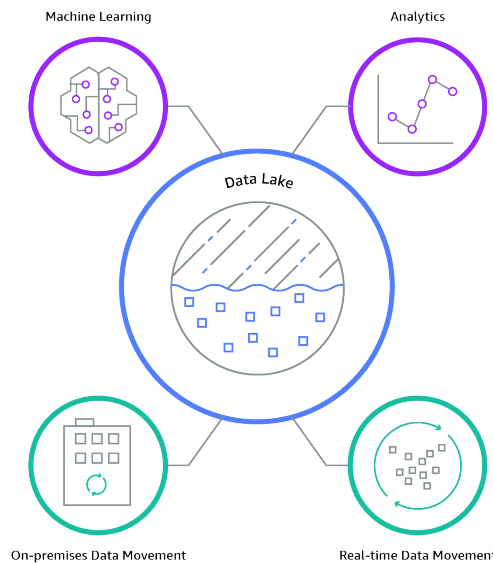
Data Lake

AWS S3(Simple Storage Service)

Data Lake

What is a data lake?

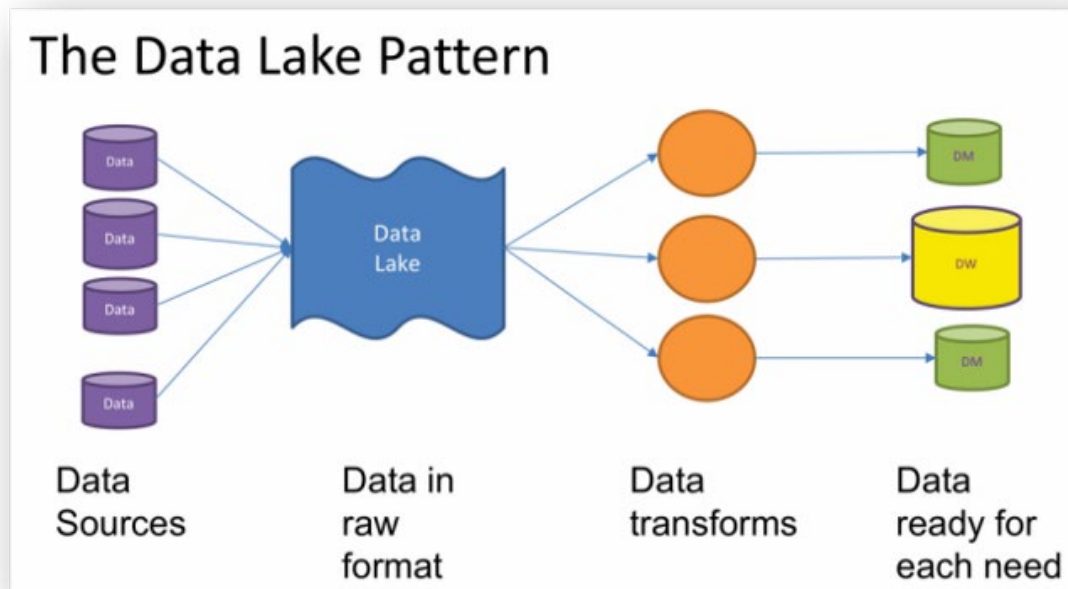
- A data lake is a centralized repository that allows you to store all your structured and unstructured data at any scale.
- You can store your data as-is, without having to first structure the data, and run different types of analytics—from dashboards and visualizations to big data processing, real-time analytics, and machine learning to guide better decisions.



On-premises Data Movement

Real-time Data Movement

© Informatica. Proprietary and Confidential.



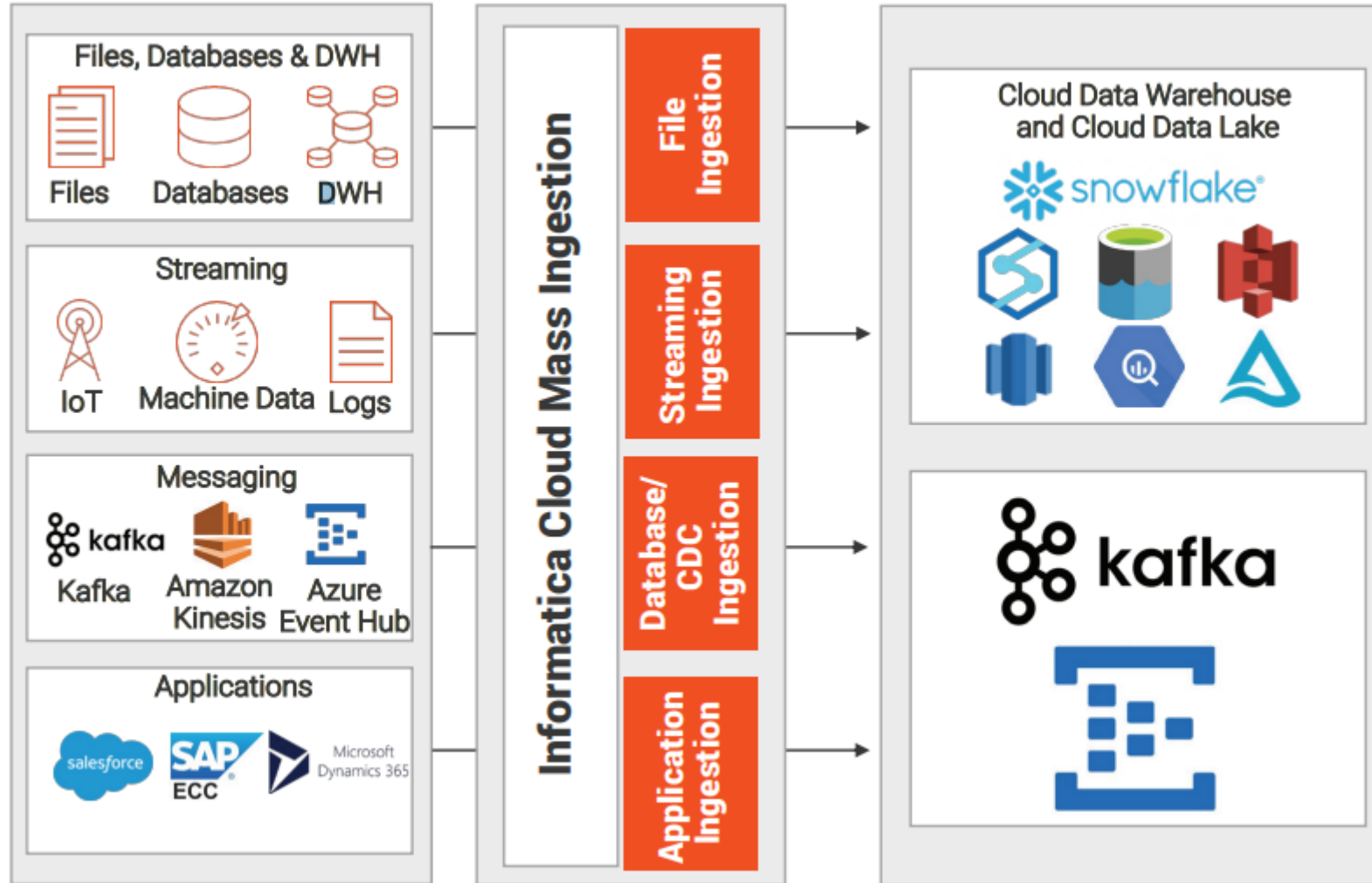
Mass Ingestion Application

Mass Ingestion Application

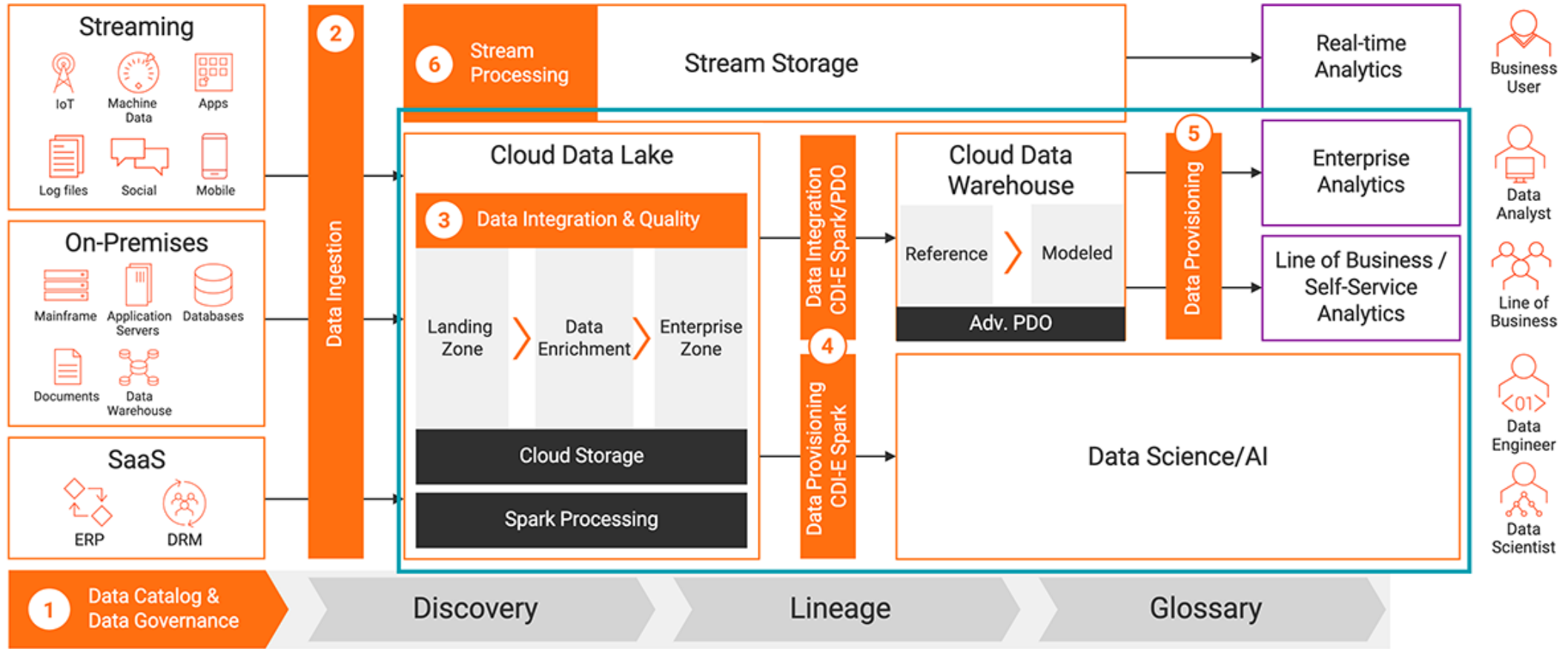
Cloud Mass Ingestion Application

The new Cloud Mass Ingestion Application capabilities enable you to ingest data from SaaS and on-premises applications—such as Salesforce, SAP ECC, and Microsoft Dynamics 365—into a cloud data warehouse or data lake such as Snowflake Data Cloud, Microsoft Azure Synapse Analytics, ADLS Gen2, Amazon Redshift, or AWS S3, as well as messaging systems such as Kafka. In addition, Cloud Mass Ingestion Application also supports schema drift and real-time monitoring of ingestion jobs with lifecycle management and alerting capabilities

CMI Services



Blueprint for a Modern Data Warehouse and Data Lake



Support Source Mass Ingestion Application

Source type	Supported load operations
Adobe Analytics	Initial load, incremental load, and combined initial and incremental load
Google Analytics	Initial load, incremental load, and combined initial and incremental load
Marketo	Initial load, incremental load, and combined initial and incremental load
Microsoft Dynamics 365	Initial load, incremental load, and combined initial and incremental load
NetSuite	Initial load, incremental load, and combined initial and incremental load
Oracle Fusion Cloud Applications	Initial load

Salesforce	Initial load, incremental load, and combined initial and incremental load
Salesforce Marketing Cloud	Initial load
SAP ERP Central Component (SAP ECC)	Initial load, incremental load, and combined initial and incremental load depending on the connection type.
SAP S4/HANA	Initial load, incremental load, and combined initial and incremental load
ServiceNow	Initial load, incremental load, and combined initial and incremental load
Workday	Initial load, incremental load, and combined initial and incremental load
Zendesk	Initial load, incremental load, and combined initial and incremental load

Support Target Mass Ingestion Application

Source type	Supported target type
Adobe Analytics	Amazon Redshift, Amazon S3, Apache Kafka, Google BigQuery, Google Cloud Storage, Microsoft Azure Data Lake Storage Gen2, Microsoft Azure Synapse Analytics, Oracle, and Snowflake
Google Analytics	Amazon Redshift, Amazon S3, Apache Kafka, Databricks Delta, Google BigQuery, Google Cloud Storage, Microsoft Azure Data Lake Storage Gen2, Microsoft Azure Synapse Analytics, Oracle, and Snowflake
Marketo	Amazon Redshift, Amazon S3, Databricks Delta, Google BigQuery, Google Cloud Storage, Microsoft Azure Data Lake Storage Gen2, Microsoft Azure Synapse Analytics, Oracle, and Snowflake
Microsoft Dynamics 365	Amazon Redshift, Amazon S3, Apache Kafka, Databricks Delta, Google BigQuery, Google Cloud Storage, Microsoft Azure Data Lake Storage Gen2, Microsoft Azure Synapse Analytics, Oracle, and Snowflake
NetSuite	Amazon Redshift, Amazon S3, Apache Kafka, Databricks Delta, Google BigQuery, Google Cloud Storage, Microsoft Azure Data Lake Storage Gen2, Microsoft Azure Synapse Analytics, Oracle, and Snowflake

Support Target Mass Ingestion Application

Oracle Fusion Cloud	Amazon Redshift, Amazon S3, Apache Kafka, Databricks Delta, Google BigQuery, Google Cloud Storage, Microsoft Azure Data Lake Storage Gen2, Microsoft Azure Synapse Analytics, Oracle, and Snowflake	SAP	Using SAP ODP Extractor connector- Amazon Redshift, Amazon S3, Apache Kafka, Databricks Delta, Google BigQuery, Google Cloud Storage, Microsoft Azure Data Lake Storage Gen2, Microsoft Azure Synapse Analytics, Oracle, and Snowflake Using SAP Mass Ingestion connector- Snowflake
Salesforce	Amazon Redshift, Amazon S3, Apache Kafka, Databricks Delta, Google BigQuery, Google Cloud Storage, Microsoft Azure Data Lake Storage Gen2, Microsoft Azure Synapse Analytics, Oracle, Microsoft Azure SQL Database, and Snowflake	ServiceNow	Amazon Redshift, Apache Kafka, Databricks Delta, Amazon S3, Google BigQuery, Google Cloud Storage, Microsoft Azure Data Lake Storage Gen2, Microsoft Azure Synapse Analytics, Oracle, and Snowflake
Salesforce Marketing Cloud	Amazon Redshift, Amazon S3, Databricks Delta,	Workday	Amazon Redshift, Amazon S3, Databricks Delta, Google BigQuery, Google Cloud Storage, Microsoft Azure Data Lake Storage Gen2, Microsoft Azure Synapse Analytics, Oracle, and Snowflake
		Zendesk	Amazon Redshift, Amazon S3, Apache Kafka, Google BigQuery, Google Cloud Storage, Microsoft Azure Data Lake Storage Gen2, Microsoft Azure Synapse Analytics, Oracle, and Snowflake

DEMO

Cloud Mass Ingestion Application
walkthrough and execution



Thank you

Email: chak@informatica.com

