

Connecting Informatica PC or Cloud to Amazon S3

Introduction

Amazon S3 is currently one of the best cloud storage options. We can connect to Amazon S3 bucket with various ETL tools. We can transform the data into files and place in Amazon S3. This document helps us to know, how the data transforms from different databases or files to Amazon S3 bucket through Informatica PowerCenter (PC) or Informatica Cloud.

A. Informatica PC - Amazon S3

We can connect Informatica PC and Amazon S3, which is a cloud based store and can store objects and delimited files which are in string 256 format on buckets.

Steps for loading Data to Amazon S3:

- Pull the data from various sources and load it as files into the AWS S3 bucket.
- Configure the AWS S3 bucket and use the script to load the files into AWS S3 bucket.
- Make the scripts reusable and reuse them to load as many times as required without having to modifying them.

Advantages

- This process enables reduction of costs.
- Enables script reusability across various environments as the S3 locations are not hard-coded.

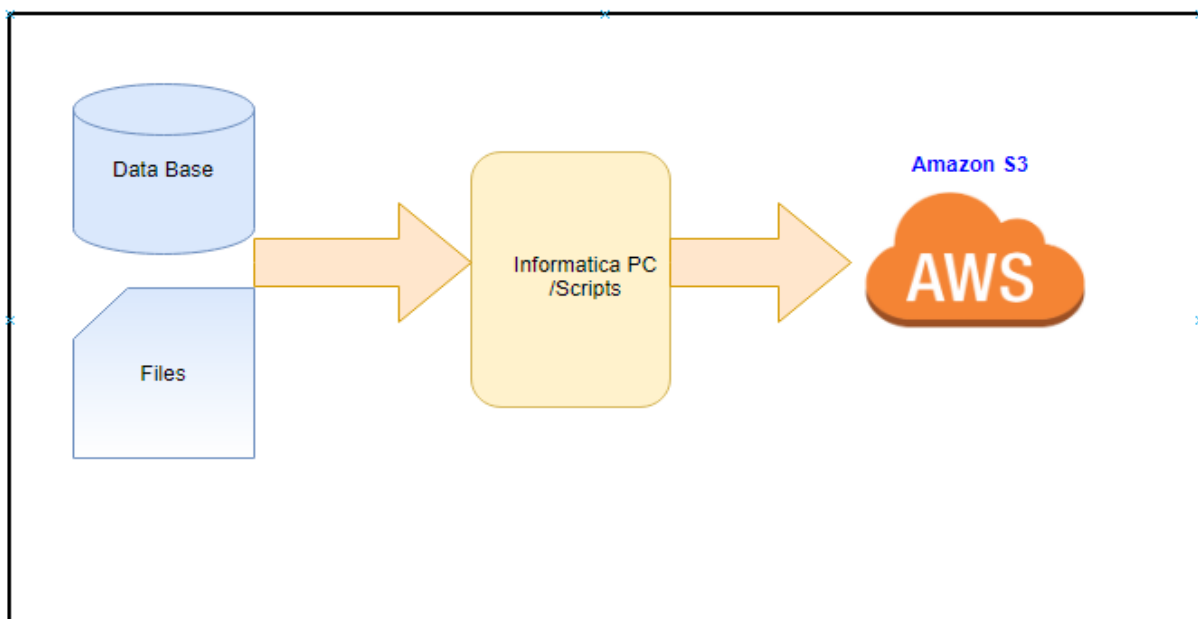


Fig: Flow - Informatica PC connecting to Amazon S3

B. Informatica Cloud - Amazon S3

When the data is available to be pulled from either database or files,

- Text file or table gets moved to source directory.
- Files are uploaded to S3 bucket.

Advantages

- This process can directly connect to Amazon S3 bucket.
- Once the connector is created, it can be reused in several jobs.

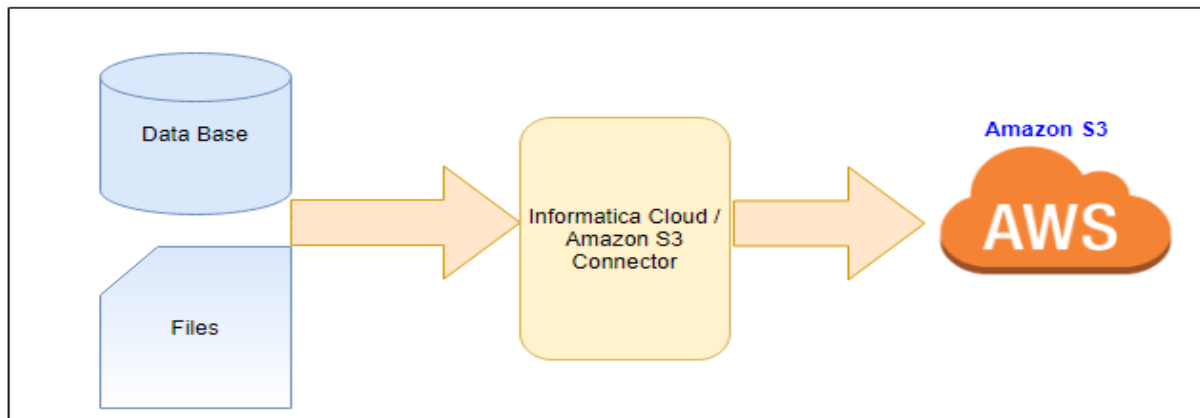
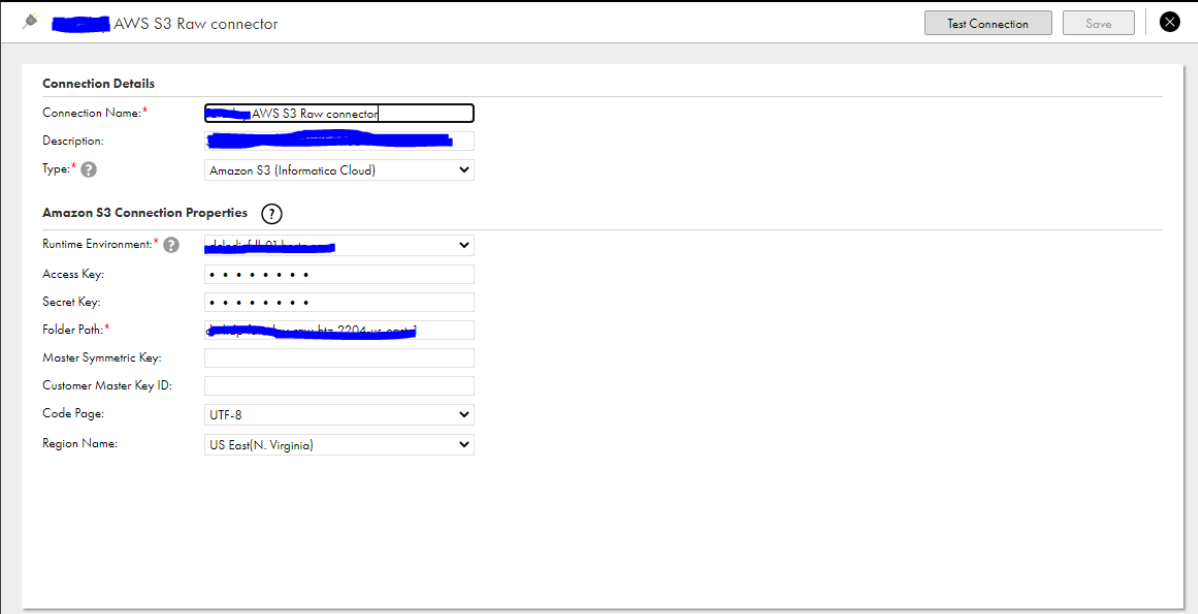


Fig: Flow - Connecting to Amazon S3 using Informatica Cloud

The below AWS signature settings are applicable to all the event sources and event targets in the connection:



The screenshot shows the configuration interface for an "AWS S3 Raw connector". The "Connection Details" section includes:

- Connection Name: AWS S3 Raw connector
- Description: [Redacted]
- Type: Amazon S3 (Informatica Cloud)

The "Amazon S3 Connection Properties" section includes:

- Runtime Environment: [Redacted]
- Access Key: [Redacted]
- Secret Key: [Redacted]
- Folder Path: [Redacted]
- Master Symmetric Key: [Redacted]
- Customer Master Key ID: [Redacted]
- Code Page: UTF-8
- Region Name: US East(N. Virginia)

Buttons for "Test Connection", "Save", and a close icon are visible at the top right of the configuration window.

Screenshot depicting the setup for AWS Connector

Connecting Informatica PC or Cloud to Amazon S3

Data format of Amazon S3 Files

- Text Qualifier : Double quote ("")
- Delimiter: Colon (:)
- Escape Character : Back Slash (\)

Optimizations

We can optimize the performance while uploading data to Amazon S3. The key point to be remembered here is, before uploading the data to S3 bucket, it needs to be compressed and encrypted.

Contact for further details



Praveen Kumar Peruca

Technical Architect - Analytics Data Engineering

praveenp.in@mouritech.com

MOURI Tech