Informatica for Amazon Redshift

Benefits

- Accelerate your cloud analytics project with Amazon Redshift and Informatica Intelligent Cloud Services℠
- Rapidly Integrate Amazon Redshift with all of your multi-cloud and on-premises data sources with hundreds of out-of-the-box connectors and data transformations
- Improve productivity for developers and citizen integrators with easy-to-use GUI-driven tools and out of the box templates and wizards
- Seamlessly scale your project with horizontal and vertical scaling and automated parallel loading to Amazon Redshift via Amazon S3

Jumpstart Your Cloud Analytics Modernization with Amazon Redshift and Informatica®

Data is emanating from every device and application that we interact with. The growing abundance and variety of data available to organizations today holds the promise of maximizing every aspect of the business, from customer acquisition and retention to better operational performance. But unleashing the power of data for business success relies on your ability to rapidly combine and analyze massive volumes of disparate data spread across multi-cloud, SaaS applications and on-premises systems in meaningful ways.

Traditional approaches for building and evolving on-premises data warehouses are robust, but can take months to implement and cost millions of dollars in total cost of ownership. Amazon Redshift and Informatica enable you to rapidly and cost-effectively set up and evolve a cloud data warehouse, connect to any cloud and on-premises data source, and deliver the data-driven agility required for business success today.

Amazon Redshift provides a petabyte-scale cloud data warehousing service that's fully managed, fast, and cost-effective and that serves as a powerful launch pad for analytics applications such as Tableau, Birst and Microstrategy. It changes the dynamics of data warehousing by making it easy to provision nodes, scale on demand, and query datasets securely.

Clusters can be resized dynamically without downtime, and distributing workloads across compute nodes can optimize I/O time. Amazon Redshift's architecture leverages massive parallel processing (MPP) capabilities with columnar storage and data compression to enable timely execution of even the most complex queries and the resulting business insights.

However, a data warehouse is only as good as the data held within it. In order to fully realize the benefits of Amazon Redshift, you must rapidly load it with trusted, integrated, relevant, timely, and enriched data. Informatica Intelligent Cloud Services is a multi-tenant, fully-managed integration platform as a service (iPaaS), which can be rapidly and cost-effectively consumed on a subscription basis. It enables you to access, aggregate, synthesize, and, in parallel, load data into Amazon Redshift from numerous locations, including cloud sources such as SaaS, social and IoT, as well as on-premise systems, such as Netezza, Teradata, NOSQL and relational databases. It gives you the agility to rapidly kickoff a small cloud analytics Redshift project and seamlessly scale it up or down as data volume and needs change. Combining an enterprise-grade cloud data integration and management solution with Amazon Redshift enhances and expedites your analytics initiative and allows you to realize the benefits of Amazon Redshift to modernize your analytics.

Example of a typical enterprise data warehouse data workflow with slowly changing dimensions – where changes are tracked in different records
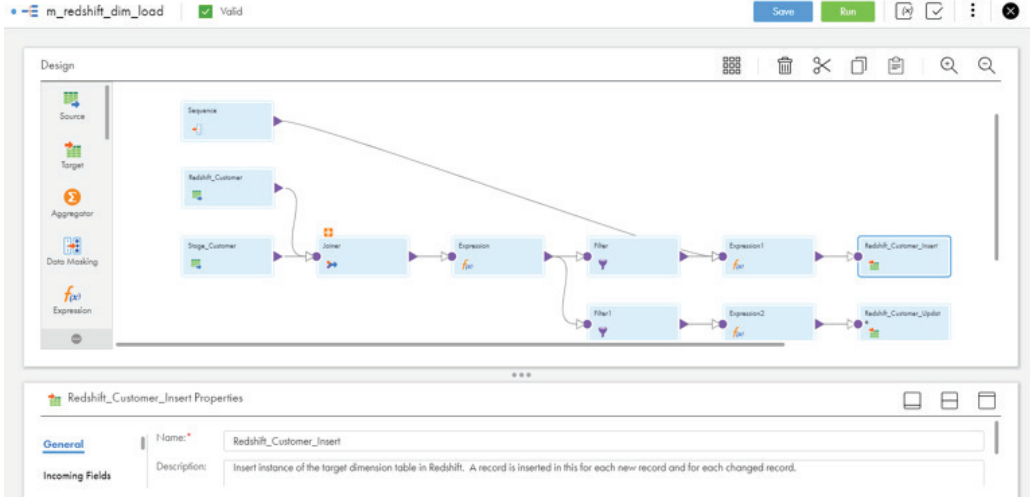
Key Features

Informatica Intelligent Cloud Services provides native, high volume, high-performance data integration with Amazon Redshift and supports out-of-the-box connectors to any cloud and on-premises data system and application, including Amazon S3, Aurora, RDS and DynamoDB. It supports secure data movement between your on-premises environment and AWS using a Secure Agent: a light-weight self-managed binary that runs in an Amazon EC2 environment, or on-premises behind your firewall, to securely access your organization's data from any location. For data that's residing in a cloud environment, Informatica Intelligent Cloud Services can use Informatica's own hosted runtime to perform the data movement. The Informatica Intelligent Cloud Services located in and Informatica- managed environment hosted on AWS enables you to develop, run and automate data integration tasks, also known as mappings, data synchronization

tasks, cloud application integration, create task flows and scheduling. Informatica Intelligent Cloud Services integration is a visual, easy-to-use metadata-driven solution, enabling self-documenting code, improved data visibility via metadata, extensive reuse in development and automation in deployment.

Informatica for AWS expedites time to production by greatly improving your ability to integrate, develop and deploy.

Integrate
Manually hand coding extract, transform and load (ETL) integrations of each data source with Amazon Redshift can slow your team down and unnecessarily tax your precious development resources. Especially with new data types and sources constantly on the rise and changing, maintaining this expertise in-house is not realistic, or productive and doesn't scale. Furthermore, manual integration tasks are not reusable across different data sources and hand coding to different environments may require different skill sets. Informatica Intelligent Cloud Services' architecture abstracts complexity and fundamentally separates integration logic from connectivity logic, so that your developers can focus on assembling the right data your business needs, without burdening them with native knowledge of the underlying systems. Furthermore, integration logic can easily be reused across different data sources. The ease of connectivity, reuse and highly productive developer environment greatly enhances the productivity of your team and allows them to deliver the right data in a timely manner into your Amazon Redshift cluster.

Develop
Informatica Intelligent Cloud Services is a visual, metadata-driven, web-based development environment, designed to increase productivity of both technical developers and line-of-business 'citizen integrators', such as data analysts and end business users. The visual Cloud Designer allows developers to quickly build complex data integration mappings and synchronization between multi-cloud, on-premises data sources and Amazon Redshift. Cloud Designer transformations palette includes out-of-the-box templates such as aggregators, joiners, lookups and sensitive data masking as well as support for dynamic mappings. For line of business users, who expect self-service data integration, Informatica Intelligent Cloud Services offers easy to use point & click Data Wizards for the most common data integration and synchronization tasks, with pull-down menus for intuitive access to data sources and basic data transformations.

If you have data in files in a format that matches your Redshift tables, you can use Informatica's Mass Ingestion service to upload and copy a large number of files to the corresponding tables in Redshift using a Managed File Transfer (MFT) functionality. This is particularly useful when there are no transformations required over the data.
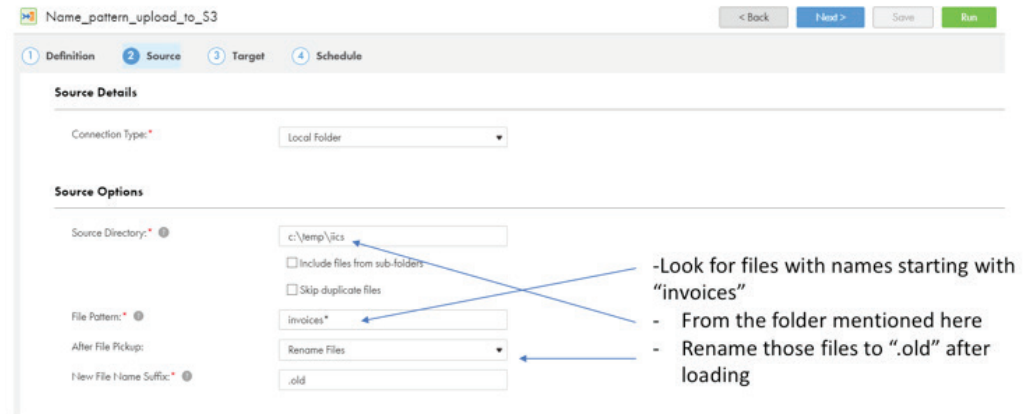
Mass Ingestion – selecting source files to upload to AWS

Deploy

Informatica Intelligent Cloud Services automates the deployment and scheduling of integration tasks and task flows with Amazon Redshift. As your project matures and you need to process increasing data volumes, you can seamlessly expand your Amazon Redshift cluster and scale your Informatica environment by clustering Informatica secure agents running on AWS EC2. Informatica Intelligent Cloud Services also supports pushdown optimization into Redshift, leveraging Redshift processing power to further increase performance with scale. In order to fully optimize Redshift performance, your deployment must harness parallel loading into your Amazon Redshift cluster via Amazon S3 staging areas. Informatica Intelligent Cloud Services further unlocks Redshift processing power by automating parallel processing tasks into Amazon Redshift, utilizing Amazon S3 for interim staging. To gain full operational confidence in production environments, Informatica Intelligent Cloud Services enables you to manage, monitor and trouble-shoot your live integration processes, managing your secure agents in multiple locations, thus giving you the confidence to catch and correct production issues early.

https://www.informatica.com/aws