Data-Driven Transformation on AWS: A Blueprint for Success

Key Benefits

- Accelerate data warehouse and data lake modernization with a metadata-driven, AI-powered integration platform as a service

- Successfully deploy a new data warehouse, data lake, or data lakehouse on AWS with high-performance data integration that connects to all data, and seamlessly integrates high volumes of data for any analytics workload

- Reduce the cost of data warehouse and data lake consolidation and modernization by leveraging the flexibility and scalability of AWS

- Ensure better business outcomes by addressing data quality problems before moving your data to Amazon Redshift and Amazon S3, and then migrate with confidence and speed

- De-risk consolidation and modernization initiatives by avoiding the challenges that come with using hand coding and multiple limited point solutions to address data management issues

- Integrate all your data with hundreds of pre-built connectors for cloud and on-premises sources, including native high-performance AWS connectors

Modernize Cloud Analytics Quickly and Cost-Effectively With Intelligent, Automated Cloud Lakehouse Data Management From Informatica

The challenge for today's enterprises is to unleash the power of their data, while getting more out of their IT investment. In order to do so, companies are moving their analytics workloads to the cloud and modernizing their applications, with the goal of increasing efficiency, productivity, and cost savings. Given that over 50% of data often comes from sources outside the company, many organizations struggle with a lack of control over data format and data quality. In addition, legacy on-premises systems are often hobbled by inefficient processes and lack the agility to quickly deliver the time-sensitive data insights required to keep companies competitive.

To overcome these hurdles, organizations are investing in cloud data warehouses, data lakes, and data lakehouses. But without the capability to automatically discover, cleanse, integrate, manage, govern, and secure data across your environment, these initiatives may fail. One-off solutions such as hand coding increase manual processes, consume resources needed by the business, and can erode trust. To provide high-quality, actionable data to the business quickly, you need an automated cloud lakehouse data management solution that enables you to have a complete view of where all your critical data resides across different silos, applications, and regions.

Informatica® on Amazon Web Services (AWS) empowers organizations to evolve their business with data-driven insights, easily scale their environment, and free up IT resources to rapidly deliver business differentiation. With Informatica's metadata-driven, intelligent cloud data management capabilities, you can realize the promise of cloud data warehouses, data lakes, and data lakehouses on AWS by automating the delivery of trusted, accurate data that drives faster innovation. Whether you are taking a cloud-first approach and building a new cloud data warehouse, data lake, or data lakehouse, or migrating your existing on-premises systems to Amazon Redshift or Amazon Simple Storage Service (Amazon S3), Informatica increases agility, reduces project risk, and provides the data quality, governance, and transparency required to meet your specific business needs.

Migrate and Modernize Your Enterprise Data Warehouse With Confidence to AWS

With the proper tools, migrating and modernizing your legacy, on-premises enterprise data warehouse (EDW) to the cloud can help your organization become truly data-driven. Informatica's AI-powered Enterprise Data Catalog (EDC) enables you to quickly discover, inventory, and organize data assets and provides a unified view of enterprise metadata to add context to your data.

To help you move assets into a new cloud data warehouse on Amazon Redshift, Informatica's complementary utility, the Intelligent Data Migration Assessment (IDMA), leverages the EDC to assess the process of migrating on-premises Teradata, Oracle, or Netezza EDWs to Amazon Redshift, a cloud data warehouse. IDMA provides an executive overview and a detailed report of the physical and business characteristics of your on-premises EDW—including the impact of sensitive data such as PII—as well as an estimate of your potential savings.

Once you identify the data you want to move and its dependencies and linkages, you can simplify integration and rapidly move large data volumes to AWS, using Informatica's cloud-native integration platform as a service (iPaaS) for ETL/ELT patterns and leveraging 150+ pre-built codeless connectors such as Amazon Relational Database Service (Amazon RDS) and Amazon Redshift. You can then utilize Informatica Intelligent Cloud Services℠ (IICS) to integrate, manage, and ensure the quality of data you want to move to AWS.

This blueprint allows you to rapidly experiment and migrate or extend your on-premises EDW to Amazon Redshift. It also enables you to protect your investment while realizing the benefits of a modern, petabyte-scale data warehouse that supports business demands for self-service, agile analytics.

Cloud Data Lakehouses on AWS: Easy. Enterprise-grade. Powerful.

In a recent survey by analyst firm TDWI, a majority (64%) of participating organizations stated that data quality and data management issues are the leading barriers to successfully delivering cloud data warehouses and data lakes.[1] In the same survey, a clear majority (86%) responded that a systematic approach to cloud data management is important to the success of their data strategy.

With Informatica and AWS, you can implement a secure, flexible, and cost-effective data lake. Amazon Redshift extends your queries to your Amazon S3 data lake without moving or transforming data. This gives you the flexibility to store highly structured, frequently accessed data in Amazon Redshift, as well as keep exabytes of structured and unstructured data in Amazon S3, and query seamlessly across both to provide unique insights that you would not be able to obtain by querying independent datasets. Together, this creates a data lakehouse, which allows you to build analytics solutions across all your data.

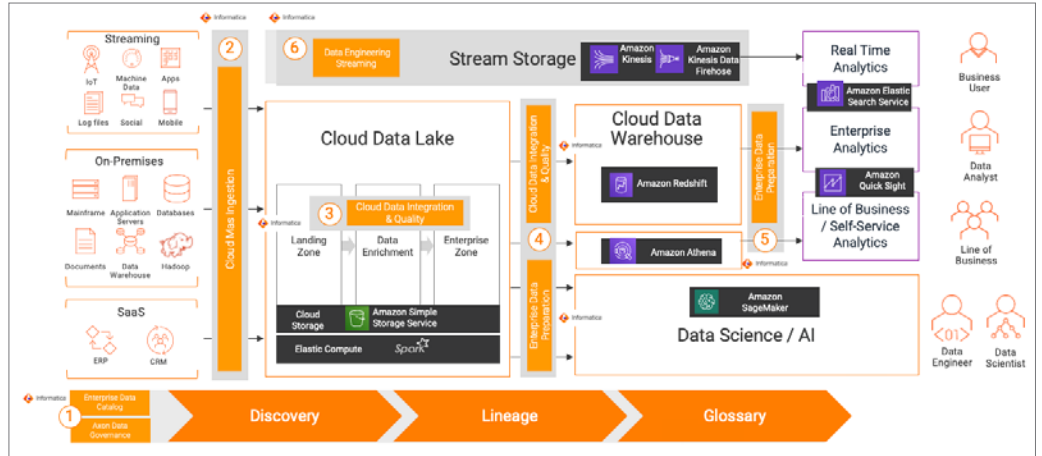[1] TDWI Best Practices Report: Cloud Data Management

Figure 1: Cloud data lake architecture: AWS ecosystem support

It is important to enable all users, from your data analysts to your data engineers, with self-service analytics to rapidly access and use data from the data lake while maintaining data governance. With Informatica's intelligent, automated cloud data lakehouse management, you can quickly and easily identify, prepare, integrate, and govern your big data pipelines to build an enterprise-scale cloud data lake. Ingest data from a variety of sources, including files, databases, change data capture, and real-time streaming data. Accelerate analytics and AI/ML projects and drive increased value from all your structured, unstructured, and semi-structured data by integrating new data sources with your existing IT investments.

Your Blueprint for a Successful Cloud Data Warehouse, Data Lake, or Data Lakehouse on AWS

With our Enterprise Data Catalog and Informatica Intelligent Cloud Services, we have the blueprint you need to build a successful data warehouse, data lake, or data lakehouse on AWS. By combining Informatica cloud lakehouse data management capabilities with Amazon Redshift, you can improve agility, innovate faster, and drive your business forward. A unified enterprise metadata management foundation enables end-to-end workstreams for data cataloging, data integration, data quality, and governance on AWS.

Data Cataloging and Metadata Management: AI-powered metadata management provides visibility across your entire data landscape with end-to-end data lineage, so you can quickly discover, inventory, and organize data assets. Informatica Enterprise Data Catalog allows you to scan and collect metadata from all your enterprise systems, curating and augmenting the metadata with business context and inferring relationships and lineage. You can use collaborative self-service data preparation to rapidly enrich, cleanse, and govern data pipelines for faster insights.

3

About Informatica

Digital transformation changes expectations: better service, faster delivery, with less cost. Businesses must transform to stay relevant and data holds the answers.

As the world's leader in Enterprise Cloud Data Management, we're prepared to help you intelligently lead—in any sector, category, or niche. Informatica provides you with the foresight to become more agile, realize new growth opportunities, or create new inventions. With 100% focus on everything data, we offer the versatility needed to succeed.

We invite you to explore all that Informatica has to offer—and unleash the power of data to drive your next intelligent disruption.

Data Integration: Informatica Intelligent Cloud Services is an enterprise-grade, cloud-native, automated, and AI-powered iPaaS. It provides data integration, critical optimization capabilities like push-down optimization, and allows for pre-built, out-of-the-box connectivity to 150+ on-premises and key AWS services such as Amazon Redshift, Amazon Redshift Spectrum, Amazon Aurora, Amazon RDS, Amazon DynamoDB, Amazon EMR, Amazon SageMaker, and Amazon QuickSight. With these capabilities, you can quickly and efficiently build data pipelines and migrate on-premises data workloads to Amazon Redshift data warehouses or Amazon S3 data lakes.

Data Quality and Governance: Intelligent, automated data quality and data governance ensure that your data is cleansed, standardized, trusted, and secure. Informatica Cloud Data Quality on AWS can help you quickly profile your data and identify, fix, and monitor data quality problems—without coding—before moving the data into your Amazon Redshift data warehouse or Amazon S3 data lake. This ensures trusted data is delivered for all your analytics initiatives.

Next Steps

Informatica products can be easily acquired and launched via the AWS Marketplace. This eliminates the need for time-consuming and costly investment in infrastructure and helps you rapidly stand up intelligent data management solutions on AWS, unleashing the power of trusted, actionable data in a cost-effective manner.

To learn more about how Informatica accelerates data modernization on AWS, visit www.informatica.com/aws or contact us.

Avis Budget Group optimized its global vehicle rental business using real-time data with Informatica and AWS. Learn how.