



Modernize your analytics by migrating to Amazon Redshift



Modernize your analytics by migrating to Amazon Redshift

Data is at the center of every application, process, and business decision. In today's digital world, organizations are grappling with exponential growth in data volumes from disparate sources. This data deluge presents both challenges and opportunities for businesses to harness valuable insights and drive innovation. The [Worldwide IDC Global Datasphere Forecast, 2023–2027](#), estimated that 129 zettabytes of data would be generated in 2023, with expectations for this volume to more than double by 2027. This means that data remains a vast untapped resource and is the key to accelerate and accomplish your organizational goals.

With the advent of generative AI, data is the differentiator that enables you to move from generic applications to generative AI applications that create real value for your customers and business. However, only 24 percent of CDOs and data executives describe themselves as data-driven, and a mere 20 percent claim to have successfully developed a data culture within their organizations¹. Data systems are often sprawling, siloed, and complex, with diverse data sets spread out across data lakes, data warehouses, cloud databases, SaaS applications, algorithms and models, and on-premises systems. Many organizations are sitting on a treasure trove of data in a variety of formats, including text, tabular, images, video, audio, geospatial, and vector data but struggle to get value out of it. With so varied data sources and data formats, organizations find it difficult to get a handle on where all their data sits, how to connect and act on that data effectively, and how to manage access to that data. And as data volumes grow, these challenges keep on intensifying to higher scales.

To truly unlock the value of your data to drive timely insights and innovation, you need to implement an end-to-end data foundation that makes it easier to work with data at every step of the data journey, for everyone who needs it in your organization. An end-to-end data strategy combines tools, resources, and processes for ingesting, storing, and querying data, analyzing data and building and working with AI and machine learning models, including large language models (LLMs) and other foundation models (FMs), and ultimately helping end users develop data-driven insights.

Only AWS provides you with the capabilities you need for an end-to-end data foundation that will serve you now and in the future. AWS works with more than 2 million customers to solve some of the most complex data problems in the world: whether it's helping BMW make data-driven decisions across their fleet of cars, supporting NASDAQ at petabyte scale to keep up with the volatility in trading volumes, empowering Global Foundries to apply advanced analytics, machine learning and AI to uncover patterns and predict future trends, or providing infrastructure and services to help Stability AI leverage massive amounts of data for generative AI foundation models. AWS has the most comprehensive set of services and capabilities to store, query, analyze and act on data at optimal price/performance, speed, flexibility and capabilities, so you can always leverage the right tool for the job. And, [Amazon Redshift](#) which is a fully managed, petabyte-scale cloud data warehouse service, lies at the core of AWS's data foundation.

Based on a series of in-depth interviews with organizations using Amazon Redshift, [IDC projects](#) that organizations achieve 503% 3-year ROI and an average of 10-month payback period. Amazon Redshift delivers direct and measurable business benefits, enabling 66% faster analytical queries and 61% productivity gains for analytics teams. Tens of thousands of customers rely on Amazon Redshift every day to enable high performance analytics workloads and deliver actionable insights for their businesses. [Amazon Redshift](#) is a fully managed, AI-powered cloud data warehouse that securely unifies all your data across databases, data warehouses, data lakes, open table

formats, streaming data services, through easy-to-use methods of data integration such as Zero-ETL, federated querying, streaming ingestion to power and personalize your business applications. With a massively parallel processing (MPP) architecture, and built-in security and compliance features, Amazon Redshift enables you to run use cases from near real-time analytics to building AI/ML/generative-AI applications, complex data processing, business intelligence dashboards and reports, and more based on both SQL and Apache Spark.

Customers modernize their data warehouses with Amazon Redshift in order to:



Achieve unmatched price-performance at any scale:

With Amazon Redshift, you can gain up to [6x better price-performance](#) compared to any other cloud data warehouse. AI-driven scaling and optimizations in Redshift learn your workload patterns based on query complexity, frequency, data volumes and continuously adjust resources to pro-actively scale to meet your price-performance objectives. When [Lexis Nexis](#) needed to adopt a modern cloud data warehouse to reduce costs and improve performance, they evaluated and ran proofs of concepts to find out Redshift was the most cost effective and highest performing solution for its data warehouse. Additionally, LexisNexis reduced the cost of the migration with the [AWS Migration Acceleration Program \(MAP\)](#), which helps companies accelerate their cloud migration and modernization, with automated migration tools, trainings, and service credits.



In terms of price and performance, Amazon Redshift scored higher than its competitors.”

Deepak Senthilkumar

Director, Software Engineering, LexisNexis



Unify all your data with zero-ETL:

Amazon Redshift helps you easily access, store, or ingest data across your data lakes, databases, data warehouses, streaming data sources, and open table formats – with no-code/low-code [zero-ETL](#) for integrated analytics, breaking data silos in your organization and eliminating the need to set up cumbersome or error prone ETL pipelines. Redshift integrates deeply across AWS services like [Amazon S3 data lake](#) for data lake foundation; operational databases such as [Amazon Aurora](#) and RDS; third-party datasets with AWS Data Exchange; streaming data services with [Amazon Redshift Streaming Ingestion](#), enabling you to access data in place, or easily ingest the data into the warehouse without custom pipelines or complex ETL.



Maximize value with comprehensive analytics and machine learning:

With Amazon Redshift's powerful analytics capabilities, you can run SQL queries, power dashboards, leverage near real-time analytics, AI/ML/generative AI applications, and build custom, rich analytics applications with SQL, Apache Spark or Python. [Redshift](#) empowers all types of data users within organizations to execute exploratory to complex analytics and [machine learning](#) accessing table formats and languages of choice without administrative overhead. With no infrastructure setup, all kinds of data users can start querying right away in [Amazon Query Editor](#).



We needed a solution that would allow us to analyze the migration process rapidly, enabling us to make quick, data-driven decisions. To address this, we turned to Amazon Aurora zero-ETL integration with Amazon Redshift to streamline our data ingestion process and eliminate the need for complex engineering work. With zero-ETL, we were able to get quick insights to drive critical technical and business decisions, saving months of effort. Also, it helped us explore new patterns at Intuit for large-scale data migrations and near real-time analytics."

Aruna Ghiware

Product Development Manager, Intuit



Our migration from legacy on-prem platform to Amazon Redshift allows us to ingest data 88% faster, query data 3x faster, and load daily data to the cloud 6x faster. Amazon Redshift enabled us to optimize performance, availability, and reliability—significantly easing operational complexity, while increasing the velocity of our end-users' decision-making experience."

Sunil Narayana

Senior Director, Analytics, Global Foundries



Innovate faster with secure data collaboration:

Share, collaborate on, and monetize data and analytics easily and securely across your organizations, regions, and even third-party data sets with no data movement or copying and with fine grained [governance, security, and compliance](#). With [Redshift Data Sharing](#), you can read from and write to databases in other Redshift warehouses in just a few clicks, enabling multiple teams to update shared data sets and collaborate on up-to-date data.



By migrating to AWS, we increased performance, increased security, and decreased latency, which has helped to increase our client base. The robust and scalable cloud computing environment that AWS offers is outstanding.”

Losif Kapoulas

Chief Information Officer, FXGT.com

Migrate with a strategic and proven approach:

Customers migrate to [Amazon Redshift](#) because they view AWS as a trusted partner for all their business needs, whether they aim to optimize costs, drive innovation, or fuel growth. Migrating data to AWS allows them to rapidly access all their data in the cloud, reducing costs while modernizing applications and increasing agility. This enables greater efficiency in delivering value to their customers through generative AI applications and driving further innovation. AWS provides organizations with an end-to-end data foundation that includes a comprehensive set of services, tools and programs to scale, ability to quickly integrate data from different sources and formats, built-in intelligence to accelerate and automate tasks, with governance for secure data access. Enterprise migrations can be complex and time-consuming but programs such as the [AWS Migration Acceleration Program \(MAP\)](#) can help you accelerate your cloud migration and modernization journey with an outcome-driven methodology, using a proven three-phased framework (Assess, Mobilize, and Migrate and Modernize).

Like any other IT project, the first step in your data warehouse migration is to identify and define your migration goals. The migration goals can be aligned to your business needs such as cost optimization, accelerate innovation and growth with enhanced analytics capabilities or technical goals such as migration from legacy systems, data architecture modernization. Once you have identified your goal for data warehouse migration, you need to evaluate your migration approach.

Most migrations often fall into one of the following three approaches:

- **Lift and shift:** Move your existing data warehouse “as-is” to the new platform with minimal modifications and minimal disruption to your existing processes and applications. It is suitable for simpler data warehouses or when you have a high degree of compatibility between the source and new platforms.
- **Lift and rewrite:** As you migrate to a new data warehouse, identify opportunities to streamline processes and employ the new platform’s capabilities to achieve your business goals.
- **Rearchitected:** Completely redesign your data warehouse to fully exploit the functionalities of the new platform and achieve your migration goals.

Once you’ve set your organizational goals and defined your migration strategy and approach, it’s time to meticulously plan your data warehouse migration through various phases.

Here’s a list of all of the steps you need to follow for migration to Amazon Redshift:

Phase I: Assessment

Discovery of Workload and Integrations:

Conducting discovery and assessment is a critical step in the migration process. This phase helps identify potential challenges, assess the complexity of the migration, and gather necessary information to plan and execute the migration effectively. It also gives a high-level understanding of the costs and resources required for the migration. This includes:

- **Data Profiling and Assessment:** This step involves analyzing the schema, types of objects in the schema, object data types including complex data types such as arrays, JSON, or custom data types, table sizes, dependencies, custom user defined functions, historical data volume to be migrated, and daily incremental data volume.
- **Code and Query Assessment:** This involves assessing the compatibility of existing SQL code, including queries, stored procedures, and functions etc., to evaluate complexity of the existing processes. AWS Schema Conversion Tool (AWS SCT) can help identify any unsupported features, syntax, or functions that need to be rewritten or replaced to ensure seamless integration with Redshift.
- **Performance and Scalability Assessment:** This includes identifying performance bottlenecks, concurrency issues, or resource constraints that may hinder optimal performance.
- **Application Integrations & Mapping:** This step helps to identify and document all the applications and business processes that rely on the data warehouse to understand the data flows.
- **Security and Access Control Assessment:** This includes reviewing the existing security model, including user roles, permissions, access controls, any compliance requirements and industry regulations that need to be adhered to.

Dependency Analysis:

This step involves understanding dependencies between objects for a successful migration. You can leverage system catalog views and custom queries on on-premises data warehouses to create a comprehensive object dependency report.

Effort Estimation:

You can estimate the migration efforts through:

- **Object Classification and Complexity Assessment:** Based on the discovery findings, you can categorize objects (tables, views, stored procedures) based on the migration complexity. The migration complexity depends on the feature parity between the source and target data warehouse.
- **Migration Tools:** Utilize [AWS SCT](#) to estimate the base migration effort per object type. AWS SCT can automate schema conversion, data type mapping, and function conversion, reducing manual effort.
- **Additional Considerations:** This includes additional tasks beyond schema conversion, such as data cleansing, schema optimization for Redshift performance, unit testing of migrated objects, and migration script development for complex procedures.
- **Team Sizing and Composition:** With a clear picture of the estimated effort, you can now size the team and based on the skillsets required for various migration tasks, you assemble a team with the right expertise.

Phase II: Migration wave planning

Migration waves can be carved out in multiple ways as given below:

- **Dependency-based Wave Delineation:** Objects with no or minimal dependencies will be prioritized for earlier migration waves, while those with complex dependencies will be migrated in subsequent waves.
- **Logical Schema and Business Area Alignment:** This allows to migrate related data objects together, minimizing disruption to specific business functions.

Phase III: Migration data warehouse

- **Code Conversion:** You can use the [AWS SCT](#) as the first step in the code conversion journey. AWS SCT is a powerful tool that can streamline the database schema and code migrations to Amazon Redshift.
- **Data Validation:** You can leverage AWS native tools such as [AWS Glue Data Quality](#) or commercial data validation tools like [Collibra Data Quality](#), scripting languages like Python, SQL, or Bash to develop custom data validation scripts tailored to your specific data validation requirements or utilize open-source data validation tools like [Amazon Deequ](#) or [Great Expectations](#).

Phase IV: Measure and benchmark key performance KPIs

For multi-tenant Redshift implementation, key performance indicators are segmented at platform level, tenant level, and consumption tools level.

Phase V: Monitoring Redshift performance and continual optimization

Monitor and measure Redshift performance against target KPIs at each phase and implement continual optimization. As new workloads are on boarded at each phase of the migration, it is recommended to perform regular Redshift cluster reviews and analyze query pattern and performance.

Get started with Amazon Redshift today

If you still have some questions,

- Read and bookmark this [blog](#) outlining how to plan and execute a successful migration to Amazon Redshift.
- [Schedule a consultation](#) with an Amazon Redshift specialist to discuss your specific migration needs.

If you're ready to migrate,

- Dive deeper to understand more about [Amazon Redshift](#).
- Evaluate Amazon Redshift first hand with a [Proof-of-Concept \(POC\)](#) to experience its capabilities and suitability for your organization.
- Reach out to [Amazon Redshift Partners](#) to connect with our worldwide network of partners to guide your migration process.

With our experience of migrating thousands of enterprise customers, AWS can help you at each step of your migration journey through our end-to-end services and programs such as the [AWS Migration Acceleration Program \(MAP\)](#). MAP is a comprehensive and proven cloud migration program, with tools that reduce costs and automate and accelerate execution, tailored training approaches and content, expertise from Partners in the [AWS Partner Network](#), a global partner community, and AWS investment. Through MAP, you can build strong AWS cloud foundations, accelerate and reduce risk, and offset the initial cost of migrations. Leverage the performance, security, and reliability of the cloud. Begin your journey today.

