Service Name

Amazon Redshift

Service Overview

Amazon Redshift is a fully managed, petabyte-scale cloud data warehouse service designed for analyzing structured data using standard SQL and business intelligence tools. It uses massively parallel processing (MPP) architecture with columnar storage technology to deliver fast performance at scale. Amazon Redshift offers both provisioned clusters and serverless options, enabling organizations to run complex analytical workloads without managing infrastructure while providing the performance of traditional data warehousing engines at significantly lower costs.

Key Use Cases

- Business Intelligence and Reporting: Running enterprise BI workloads, generating dashboards, and creating analytical reports for sales, finance, and operational data
- Real-time Analytics: Analyzing streaming data from IoT devices, clickstreams, gaming telemetry, and social media for immediate insights and decision-making
- Data Lake Analytics: Querying data stored in Amazon S3 using Redshift Spectrum without loading data into the warehouse, enabling analysis of huge volumes of structured and semi-structured data
- Machine Learning and Predictive Analytics: Using Redshift ML to create, train, and deploy ML models directly within the data warehouse for fraud detection, risk scoring, and customer churn prediction
- Cross-organizational Data Sharing: Securely sharing live data between different teams, departments, or business units using data sharing capabilities without copying or moving data

Features

- **RA3 Instances with Managed Storage**: Separate compute and storage scaling with high-performance query processing and automatic storage management
- **Concurrency Scaling**: Automatically provisions additional compute capacity to handle varying workload demands without performance degradation
- **Redshift Spectrum**: Query data directly in Amazon S3 using SQL without loading it into Redshift tables
- **Zero-ETL Integrations**: Near real-time data replication from Aurora, RDS, DynamoDB, and enterprise applications without complex ETL pipelines
- **Serverless Option**: Automatically manages compute capacity with AI-driven scaling and optimization without infrastructure management

- **Data Sharing**: Secure sharing of live data between separate Redshift clusters without copying or moving data
- **Redshift ML**: Create, train, and apply machine learning models using SQL commands with SageMaker integration
- **Multi-AZ Deployment**: High availability with automatic failover and data replication across multiple Availability Zones

Value Proposition

Amazon Redshift provides up to 6x better price performance compared to other cloud data warehouses and up to 7x better performance for high concurrency workloads. It offers the flexibility of both provisioned and serverless deployment models, allowing customers to optimize for their specific needs:

- The service eliminates the complexity of traditional data warehousing with automatic scaling, built-in security features including end-to-end encryption, and seamless integration with the broader AWS ecosystem
- Customers benefit from no upfront costs, pay-as-you-go pricing, and the ability to scale from gigabytes to petabytes without performance degradation

Service Integration

Amazon Redshift integrates natively with numerous AWS services including Amazon S3 for data lake analytics through Redshift Spectrum, Amazon Aurora and RDS for zero-ETL data replication, Amazon SageMaker for machine learning capabilities, Amazon Kinesis for streaming data ingestion, AWS Glue for ETL processes, Amazon QuickSight for business intelligence dashboards, AWS IAM for security and access control, Amazon CloudWatch for monitoring, AWS Backup for centralized backup management, and Amazon VPC for network isolation. It also supports integration with AWS Lake Formation for data cataloging and governance.

Onboarding and Offboarding

Customers can start with Amazon Redshift by creating either a provisioned cluster or serverless workgroup through the AWS Management Console, CLI, or API. The Getting Started experience includes sample datasets and query examples to familiarize users with the service:

- For provisioned clusters, customers choose node types and cluster size based on their requirements
- For serverless, they simply specify base capacity settings
- The service provides automated setup of VPC configurations, security groups, and parameter groups with sensible defaults, allowing users to begin querying data within minutes of creation

Getting Started Guide: https://docs.aws.amazon.com/redshift/latest/gsg/

Data Removal

To offboard from Amazon Redshift, customers can delete their clusters or serverless workgroups, which removes all associated compute resources. Manual snapshots must be explicitly deleted to remove data permanently, while automated snapshots are deleted when the cluster is terminated. For complete data removal, customers should delete all snapshots, remove any cross-region snapshot copies, and clean up associated resources like parameter groups and subnet groups to ensure no residual data or costs remain.

Backup/Restore and Disaster Recovery

Amazon Redshift provides automated and manual snapshot capabilities with configurable retention periods. Snapshots are stored in Amazon S3 and can be copied across regions for disaster recovery. The service supports point-in-time recovery through continuous automated snapshots taken every 8 hours or after 5 GB of data changes. Multi-AZ deployments offer automatic failover with sub-minute recovery times and synchronous data replication for high availability scenarios.

Backup Capabilities

Amazon Redshift includes comprehensive backup capabilities with automated snapshots taken periodically and manual snapshots created on-demand. The service integrates with AWS Backup for centralized backup management and policy-based backup scheduling. Backup retention periods are configurable from 1 to 35 days for automated snapshots, while manual snapshots can be retained indefinitely or for specified periods up to 3,653 days.

Disaster Recovery

Amazon Redshift supports disaster recovery through cross-region snapshot copying, Multi-AZ deployments for automatic failover, and integration with AWS Backup for centralized disaster recovery management. While not directly integrated with AWS Elastic Disaster Recovery, customers can implement disaster recovery strategies using snapshot restoration, data sharing across regions, and architectural patterns that leverage multiple AWS regions for business continuity.

Recovery Objectives

Amazon Redshift helps meet RPO objectives through automated snapshots every 8 hours or after 5 GB of data changes, providing recovery points with minimal data loss. RTO objectives are supported through Multi-AZ deployments with sub-minute automatic failover, elastic resize for quick capacity changes, and snapshot restoration capabilities. Cross-region snapshot copying enables disaster recovery with RPO and RTO aligned to business requirements and snapshot frequency configurations.

Service Constraints

Service Quotas: https://docs.aws.amazon.com/redshift/latest/mgmt/amazon-redshift-limits.html

FAQ: https://aws.amazon.com/redshift/faqs/

Technical Requirements

Service Documentation: https://docs.aws.amazon.com/redshift/

User Guide: https://docs.aws.amazon.com/redshift/latest/mgmt/

API Reference: https://docs.aws.amazon.com/redshift/latest/APIReference/

Pricing Overview

Please see the AWS UK G Cloud 15 Pricing Document accompanying this service in the Digital Marketplace.

Public Pricing: https://aws.amazon.com/redshift/pricing/

Cost Optimization

Amazon Redshift offers multiple cost optimization strategies including Reserved Instances with up to 75% savings over on-demand pricing, Serverless Reservations providing up to 24% discounts on committed RPU usage, and pause/resume functionality for development environments. Customers can optimize costs through proper data compression, efficient table design, and using Redshift Spectrum for infrequently accessed data. Concurrency Scaling provides one hour free per day, and managed storage in RA3 instances prevents over-provisioning by scaling storage independently of compute resources.

Savings Considerations

Primary cost savings come from choosing appropriate pricing models (Reserved Instances vs on-demand), rightsizing clusters based on actual workload requirements, and leveraging Serverless for variable workloads to avoid idle capacity costs. Additional savings are achieved through data lifecycle management using Redshift Spectrum for cold data, optimizing query performance to reduce compute usage, and using pause/resume for non-production environments. Cross-region data transfer costs should be considered when implementing multi-region architectures, and backup storage costs can be managed by setting appropriate retention periods for manual snapshots.