



6 WEEKS

Modern data warehouse

via data lake with Amazon Redshift Jumpstart

A 6-week set of packaged professional services to help Amazon Web Services (AWS) customers prove-out, pilot, and accelerate their cloud data warehouse initiatives. The Jumpstart Bundle includes: Needs Assessment, Capacity Planning, Environment Setup and Deployment, Data Modeling and Loading, Report Migration, Training, and Knowledge Transfer

What you get

- Selection of Modern Data Warehouse via Data Lake Pilot Use Case (fits into the 6-week window)
- Implementation of agreed upon use case from ingestion to consumption in a Development environment
- Tactical Plan to Operation
- Foundations of the Data Lake Reference Architecture for your enterprise that includes Amazon Simple Storage Service (Amazon S3) & Amazon Redshift
- "NorthBay Data Lake Toolkit" (Code snippets, Accelerator Scripts, and Libraries)
- Validate/Suggest Ingestion, Transformation, and Reporting platforms and tips on identifying gaps
- Visualization/Reporting connected via Amazon QuickSight, Tableau, or Looker
- Catalog, ability to search and visualize the datasets in the Data Lake
- Strategy Consulting with stakeholders for medium and long term data platform and analytics needs
- Roadmap for subsequent use case implementations

AWS Managed Service included

- Ingestion & Storage - Amazon S3 for Raw File/Object Storage and enhanced (if applicable)
- Data Warehousing - Amazon Redshift
- Transformation - Basic transformations with Amazon EMR/Matillion
- Analytics/Visualization - Amazon QuickSight
- Orchestration - Data Pipeline/step functions
- Monitoring - Amazon CloudWatch
- Metadata & Data Lake
- Data Catalog - Amazon DynamoDB
- Exploration - Exploration/ visualization with Amazon Elasticsearch Service and Kibana
- Governance and Access - AWS Identity and Access Management (IAM) (if applicable)
- Automation - AWS CloudFormation
- Microservices - Amazon API Gateway & AWS Lambda (if applicable)

**The JumpStart package configuration is based on this specific set of AWS Cloud Tools and Data Source attributes. If changes are required in the services or data sources NorthBay will gladly consult with client to reconfigure the package, services, and pricing to their unique requirements.

***NorthBay will provide the client with preparatory tasks that need to be completed before the start of the Jumpstart to ensure it launches correctly.

Schedule

Weeks 1-3

- Gather requirements and validate with stakeholders
- Identify and characterize the data sources involved in the exercise
- Finalize ETL or Amazon EMR solution for ingesting the data into Amazon Redshift
- Work on the end results and determine how the collected data appears when viewing the results

Weeks 3-5

- Complete the technical design and leverage the NorthBay Data Lake Toolkit
- Implement the data collection (ETL / Amazon EMR)
- Iterate based on what the real data delivers
- Build BI/Portal front-end using Amazon QuickSight, Tableau, or Looker

Weeks 5-6

- Demonstrate results
- Final delivery of JumpStart solution

NorthBay, the first partner to earn the AWS Big Data Competency, has architected and successfully delivered AWS-based Data Lakes & Data Warehouses for nearly 100 companies over the past six years. This hard won experience, combined with NorthBay Toolkits and Packaged Services, produces efficient, secure, and scalable Data Lakes. NorthBay's methodology of "Teaching Old Data New Tricks TM" allows companies to quickly and gracefully evolve from their legacy on-premises Data Warehouses and Data Lakes quickly and cost effectively. This methodology delivers Data Lakes that serve as the launching point for Machine Learning, Downstream Analytics, User Self Service Reporting, and Visualization tools.