

Adanto democratizes data with cloud-based Data Lake solution

Adanto provides easy access to raw data for many departments of the Silicon Valley-based global leader in the Consulting and Professional Staffing industry, and helps institutionalize data-driven digital culture throughout the entire enterprise.

Data lakes are **enterprise-wide data management platforms** that store **disparate sources of data** in its native format, until a client queries it for **analysis**. So, rather than putting the data in a purpose-built data store, data is moved into a **Data Lake** in its **original format**. By consolidating data, the information silos are eliminated which increases **information use and sharing**. It also lowers costs through server and license reduction, cheap scalability, flexibility for use with new systems, and the ability to keep the data until the data consumer – a **programmer** or a **business user**– is ready to use it.

The Challenges:

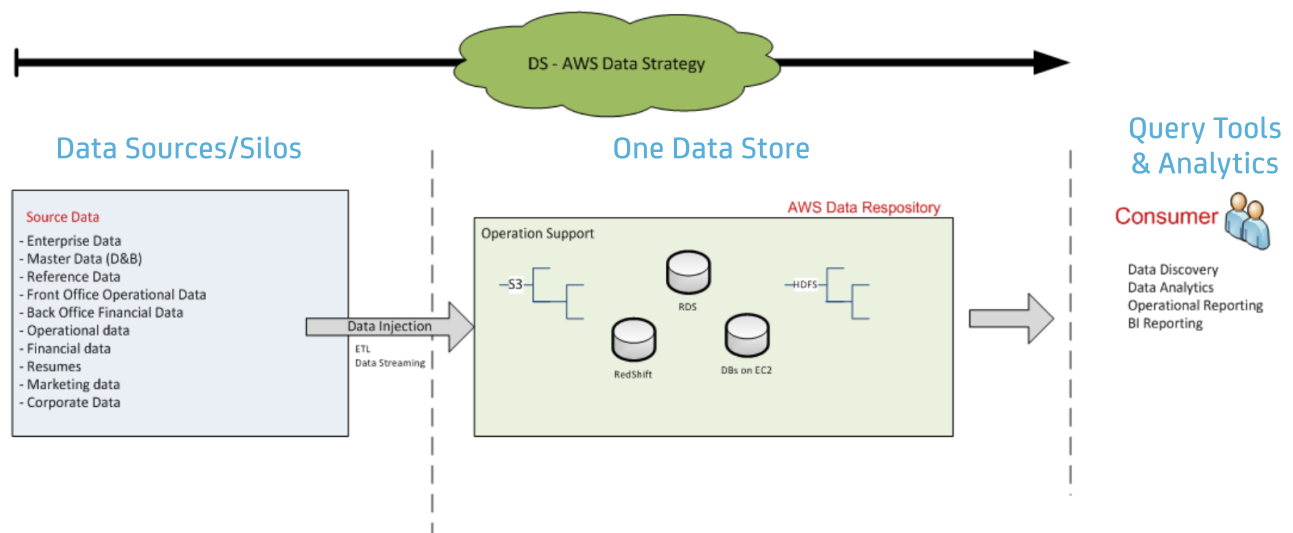
- Poor agility and accessibility for data analysis
- Data and Information silos
- Lack of information use and sharing for business decision making
- Increasing cost of multiple server and license proliferation, IT complexity
- Very expensive scalability, lack of flexibility for use with new systems

The Objectives:

- **Single store** for all the **raw data** for anyone in department to analyze
- Set of **incremental load processes**
- Data **governance procedures**
- Building thematic/departmental/business line – **central data marts**
- Building **analytic applications** for various business needs

Data lakes are **infrastructure** components supporting **systems of innovation**. Systems of innovation target creating new business models, products or services with a fail-fast mentality. However, successful innovation means making investment to scale. It is this last point, around scaling innovation, that requires a deliberate approach to designing the data lake and integrating it with your existing infrastructure to make the leap from experiments to reliable insights.

The Solution:



Data Sources/Silos:

- **>60 data sources**
- **>200 GB of** new data per day

One Data Store (Data stored in different AWS cloud-based data stores based on data type):

- **Amazon S3**
- **Amazon EC2 (Elastic Compute Cloud** service for secure, scalable compute capacity)
- **Amazon Redshift** (data warehouse for standard **SQL queries & BI tools**)
- **Amazon RDS** (relational database for many instance types: **PostgreSQL, MySQL, Oracle Server, Microsoft SQL Server**)
- **Apache Sqoop** (O/S tool for bulk data transfers)
- **Amazon HDFS (Parquet) (HadoopCluster with EMR - Elastic MapReduce)**

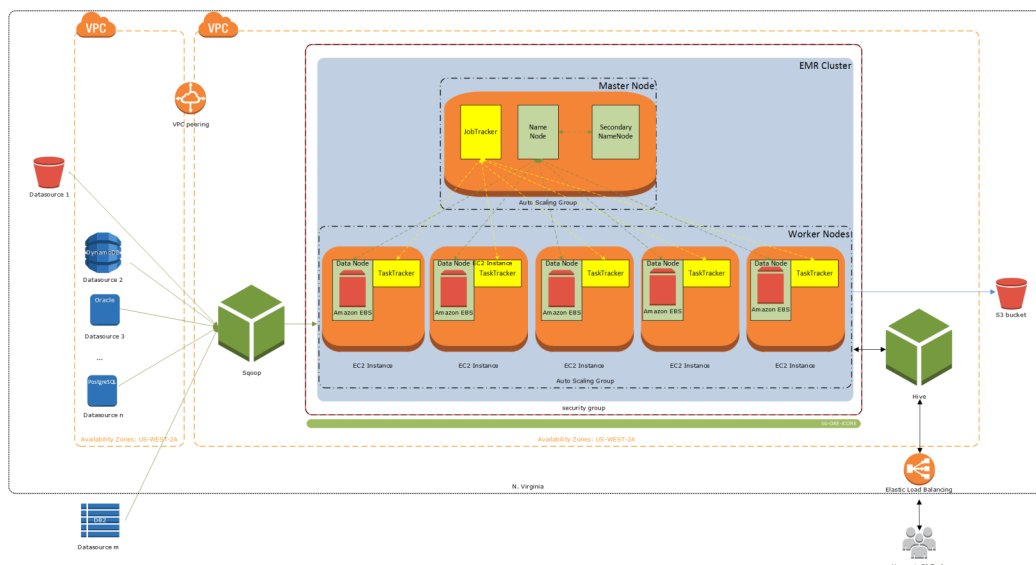
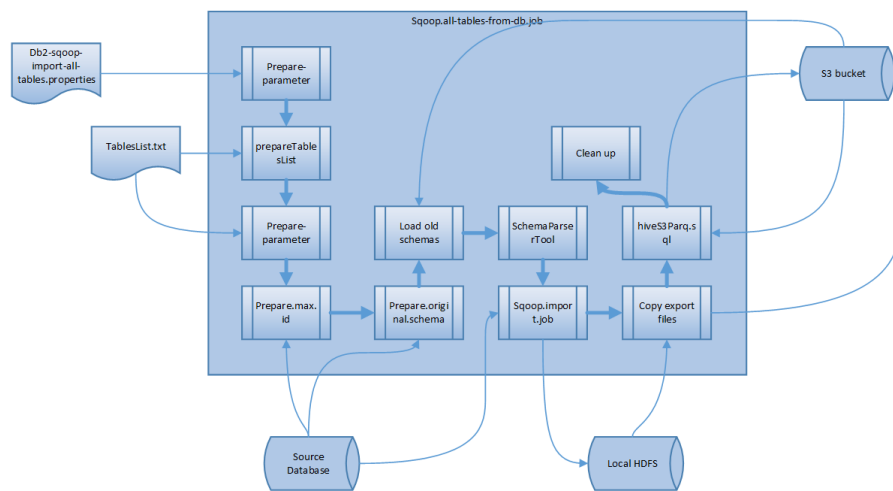
Query Tools & Analytics:

- **Apache Hive, Pig, Spark** (O/S database query interface tools to HDFS & processing engine)
- **R** (O/S statistical programming language for data mining and statistical computing)
- **Mahout/scikit-learn** (O/S tools for building Machine Learning apps)
- **Pentaho, QlikView, PowerBI, SAS** (data analytics, business intelligence and reporting tools)

Solution Features & Benefits:

- Data stored in an **inexpensive data store** - the **Amazon S3 Bucket**
- Structure built in **parquet file format** used in **HDFS/Hive** to **query data**
- Cloud-based **Hadoop/Spark** cluster set up in **AWS data center** with **autoscale** functionality.
- Incremental **load processes** run on **EMR** cluster in AWS and execute daily data pull using **Apache Sqoop**
- Unleashed the **power of business intelligence** at the business user fingertips
- Provided custom **reporting & reporting tools**
- Enabled **machine learning** efforts to uncover the **hidden potential** of the available data
- **Optimized** and **automated** the **business processes** basing on related data

High Level Data Flow Diagram



About Adanto

Adanto is an emerging global software services and solutions company, incorporated in Delaware, serving clients in the EU and the USA from its five global locations.

Combining vast experience with comprehensive capabilities across many technologies, industries and business functions, Adanto collaborates with clients to **accelerate their innovation** and help them become **high-performance digital businesses**.

www.adanto.com
+1 (888) 774-6647

adanto
s o f t w a r e

To find out more about Adanto and how to best harness Adanto Professional Services on your quest to accelerate innovation, please contact:

In USA:
Michael Kinalski
+1 208 841 6598
michael.kinalski@adanto.com

In UK:
Alexander Hill
+44 (0) 781 806 8861
alexander.hill@adanto.uk

In Poland:
Grzegorz Szutkiewicz
+48 692 455 214
grzegorz.szutkiewicz@adanto.com