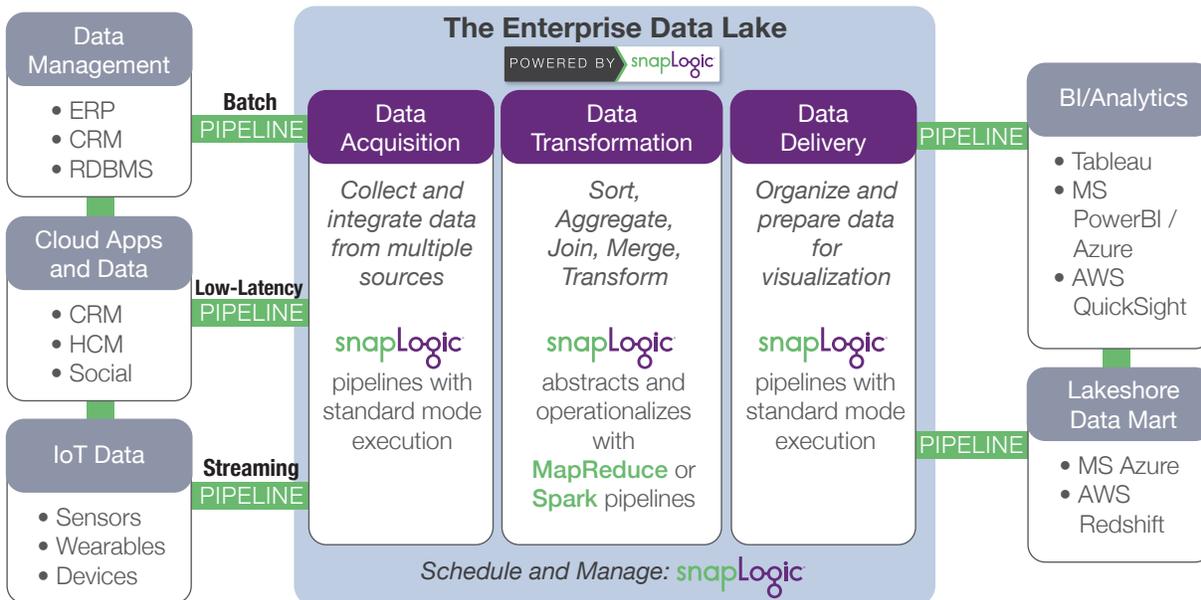




SNAPLOGIC INTEGRATION FOR BIG DATA

SnapLogic is the only unified data and application integration platform as a service (iPaaS). The SnapLogic Elastic Integration Platform has a streaming architecture that supports real-time, triggered and low-latency enterprise integration requirements plus the high volume, variety and velocity of big data integration in the same easy-to-use interface. SnapLogic's distributed architecture is a natural fit for consuming and moving large data sets residing on premises, in the cloud, or both, and delivering them to and from the data lake.

- Translate visual data flow pipelines into MapReduce or Spark jobs without scripting
- 400+ pre-built, intelligent connectors - called Snaps - for HBase, Kafka, Cassandra, MongoDB, AWS Redshift, plus hundreds of other applications and data stores
- Read data from and write to HDFS using a variety of file formats including RCFile, Parquet, ORC, AVRO, SequenceFile, and others
- Leverages YARN natively to utilize Hadoop resources to execute data pipelines
- Hybrid batch and streaming support for a range of computations including data cleansing, analytics, machine learning, as well as conventional extract, transform, and load (ETL)
- Compatible with Cloudera, Hortonworks, and Microsoft HDInsight



Powering the Enterprise Data Lake

SnapLogic accelerates development of a modern data lake through:

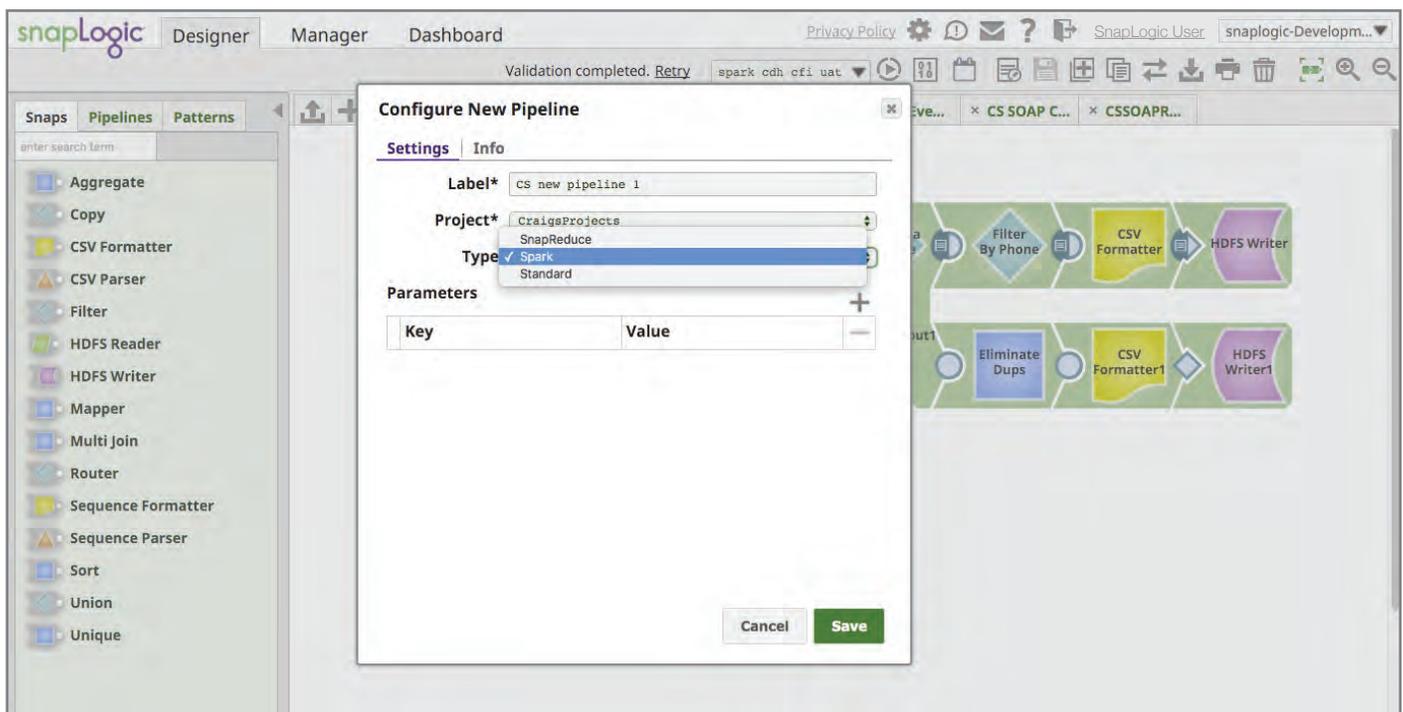
- **Data acquisition:** SnapLogic goes beyond developer tools such as Sqoop and Flume with a cloud-based visual pipeline designer for data ingestion, and pre-built connectors for 400+ structured and unstructured data sources, enterprise applications and APIs.
- **Data transformation:** Minimize the manual tasks associated with data shaping and make data scientists and analysts more efficient. SnapLogic includes Snaps for data preparation tasks such as joins and unions that eliminate scripting.
- **Data delivery:** Make data processed on Hadoop or Spark easily available to off-cluster applications and data stores such as statistical packages, analytics and business intelligence tools.

Spark and MapReduce Pipelines Without Coding

SnapLogic’s platform-agnostic approach decouples data processing specification from execution. As data volume or latency requirements change, the same pipeline can be used just by changing the target data processing mode - MapReduce, Spark, or standard mode. SnapLogic can run natively on a Hadoop cluster as a YARN-managed resource that elastically scales out to power big data analytics.

The SnapLogic visual designer helps to operationalize the MapReduce and Spark processing frameworks by enabling users to create data pipelines without coding. This puts the power of MapReduce and Spark into the hands of non-experts.

Whether MapReduce, Spark or other big data processing frameworks, SnapLogic allows customers to adapt to evolving data lake requirements without locking into a specific framework.



The SnapLogic Elastic Integration Platform

A single platform for application and data integration, SnapLogic boosts developer productivity by requiring no hand coding, no schema enforcement, and rapid development via a robust designer interface. At the same time, SnapLogic is intuitive enough to provide self-service integration capabilities for business analysts, data scientists and non-developers. With SnapLogic you eliminate the effort and maintenance involved with traditional ETL tools and remove data silos so your entire organization can connect faster. It provides:

- Fast time-to-value
- The ability to handle the most complex cloud-to-cloud and cloud-to-on premises use cases
- A dramatic boost to developer productivity
- Self-service for non-developers

