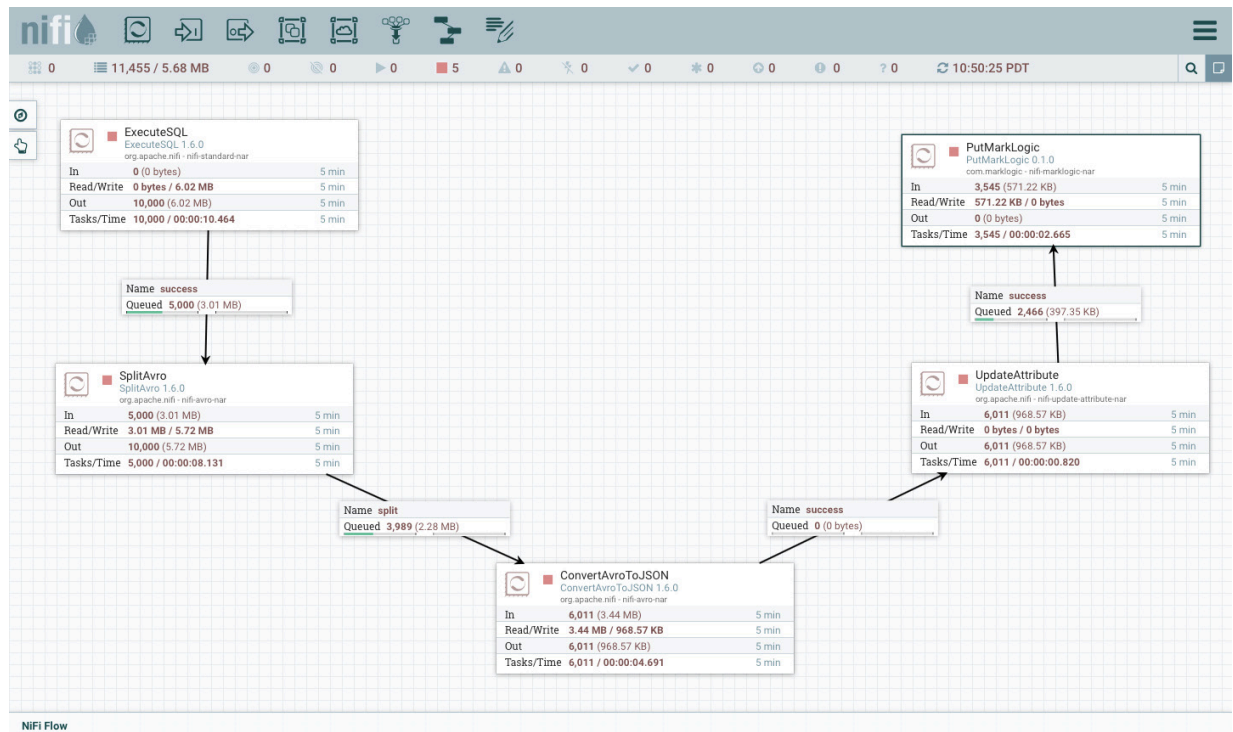


Apache NiFi

When integrating data, organizations need fast and easy ways to extract data from any source system and get it into the target system. Apache NiFi was built to automate the flow of data and it provides a nice drag and drop, configurable user interface. MarkLogic® supports its processors built for Apache NiFi, and the integration with Apache NiFi makes it a great choice for loading data into MarkLogic.



This screenshot shows a simple process to get relational data into MarkLogic. A SQL query is executed to get data out of a relational system. Then, a NiFi processor converts the resulting Avro serialization (https://en.wikipedia.org/wiki/Apache_Avro) to data formatted as JSON, and that data is put into MarkLogic.

Key Benefits

- **Developed by the NSA** – The U.S. National Security Agency spent 10 years developing NiFi before making it open source. During development, they focused on data security, provenance, and ease of use within teams.
- **Load almost any data source** – Load data into MarkLogic from relational, Hadoop, S3, and more. There are over 50 data flow processors and it is easily configurable and extensible.
- **Easy to use interface** – NiFi has a drag and drop UI that makes it easier than ever to design the data flows to load data to MarkLogic and monitor the process.



How NiFi Works With MarkLogic

Apache NiFi comes with out-of-the-box processors to create data flows from almost any system. If a processor does not exist, organizations can quickly build their own processors. It is also possible to create templates for common data flow patterns, then publish and share them for others to benefit from and collaborate further on.

Examples of Data Sources to Load to MarkLogic with NiFi

- AMQP
- Amazon SQS
- Amazon SNS
- Amazon S3
- Azure Event Hub
- Couchbase
- CSV
- DynamoDB
- Elastic Search
- Exchange Web Services
- File System
- FTP
- Google Cloud Bucket
- Grok Reader
- HBase
- HDFS
- HTTP
- Ignite Cache
- IMAP
- JMS
- JSON Path Reader
- Kafka
- Mongo DB
- MQTT
- MySQL
- Parquet
- Pop3
- Redis
- RethinkDB
- SFTP
- SNMP
- Solr
- Spark
- Splunk
- SQL
- TCP
- Windows Event Log

About MarkLogic

MarkLogic is the world's best database for integrating data from silos, providing an operational and transactional Enterprise NoSQL database platform that integrates data better, faster, with less cost. Visit www.marklogic.com for more information.

© 2018 MARKLOGIC CORPORATION. ALL RIGHTS RESERVED. This technology is protected by U.S. Patent No. 7,127,469B2, U.S. Patent No. 7,171,404B2, U.S. Patent No. 7,756,858 B2, and U.S. Patent No 7,962,474 B2. MarkLogic is a trademark or registered trademark of MarkLogic Corporation in the United States and/or other countries. All other trademarks mentioned are the property of their respective owners.

MARKLOGIC CORPORATION
999 Skyway Road, Suite 200 San Carlos, CA 94070
+1 650 655 2300 | +1 877 992 8885 | www.marklogic.com | sales@marklogic.com