



BDM Ingestion

Automates the ingest of data – at rest, streaming or both, in a secure and fast manner using Spark

PRODUCT DESCRIPTION

BDM Data Ingestion module automates the simple ingest of data for all your data sources for onward movement to any destination. The data can be at rest, streaming or both, and connectors are available for all major types of data:

- Databases: Teradata, SQL-Server, Oracle, DB2, etc.
- Files: JSON, CSV, Avro, XML
- Streaming: Kafka, Spark Structured Streaming

Connectors for non-standard sources can be created in days where required.

BUSINESS CHALLENGE

The challenges companies face in moving data is sourcing data engineers with the required skill sets and experience to write the scripts to move the data. Engineers with these skillsets are scarce. Usually engineers without the correct experience or skill set do this work and this results in data being incorrectly ingested resulting in incorrect data at the destination. The manual coding of ingest scripts will lead to code management and deployment issues as the Data Lake scales and the data grows.

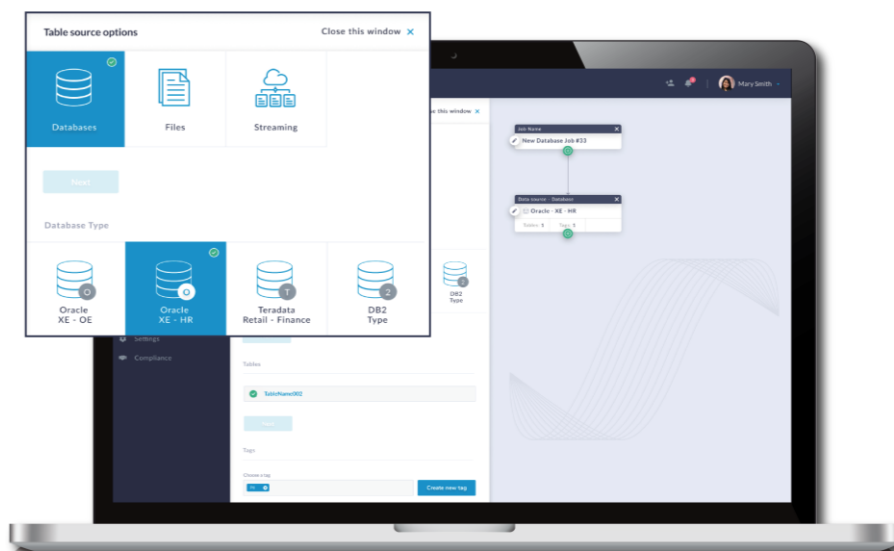
BDM solves all these issues as data can now be moved from source to destination, with no coding required.

KEY FEATURES

- All major data types and sources are supported
- Batch and Streaming options available
- Data is cleansed and normalized to ensure data is compatible
- Ingestion connectors are written and available for most data sources
- Data is cleansed and normalized to ensure data is compatible

KEY BENEFITS

- The whole process is done in memory, making it faster and more secure than other options
- It is possible to easily scale for all new data sources quickly
- Execute jobs rapidly through a simple interface
- Eliminates the need for scarce skilled engineers
- Removes the dependency on code-writing as data volumes increase



BDM SOLUTION

The BDM Ingestion module allows data to be moved from source to destination, with our automated solution eliminating coding and architecture errors.

- ✓ Spark is used as the processing environment providing enhanced security as all activities on the data are carried out in memory, with no copies being saved to disk
- ✓ Data is automatically normalized during the movement stage, where types and values are converted to enable the data to be ingested accurately into the destination source
- ✓ Schema detection is carried out on each ingest, ensuring that all changes to the source schema are immediately replicated in the destination schema
- ✓ Scalable: New data sources require a connector, which can be written in days
- ✓ All Spark code is automatically generated on the fly for each pipeline – there is no requirement to write code

TECHNICAL REQUIREMENT

Spark V 1.x or V 2.x
Atlas V 0.8 or above

APPOINTMENT

For more assistance, you can schedule a one-to-one appointment via Skype, Zoom, WebEx or phone.

FOR MORE INFORMATION.

To learn more about BDM and the Data Ingestion module
Please visit www.bluemetrix.com
| Europe +353 21 4212223 | info@bluemetrix.com

