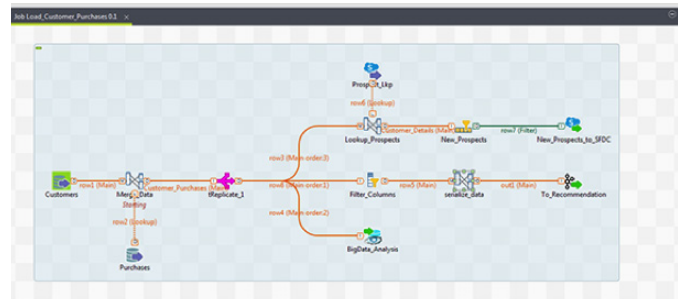


*The first data integration platform for Hadoop and Spark*

## The first data integration platform for Hadoop and Spark

Talend simplifies big data integration with graphical tools and wizards that generate native code so you can start working with Apache Hadoop, Apache Spark, and NoSQL databases today. Talend Big Data Integration platform delivers high-scale, in-memory fast data processing, as part of the Talend Data Fabric solution, so your enterprise can turn more and more data into real-time decisions.

- Blazing fast speed and scale with Spark and Hadoop
- Deliver self-service data preparation to everyone
- Optimize big data performance in the cloud
- Protect your investments with a future-proof architecture



## Blazing fast speed and scale with Spark Streaming and Hadoop

Only Talend Big Data Integration takes advantage of the massively parallel environment of Hadoop by generating native MapReduce and Spark code. Load, transform, enrich, and cleanse data inside Hadoop to leverage Hadoop's power and scale. Run 5 times faster than MapReduce using Spark in-memory data processing. Talend Studio provides access to over 900 connectors and components, plus graphical drag-and-drop tools and wizards to increase productivity and speed development.

## Drive self-service data preparation to everyone

Talend combines data preparation with big data integration in a single unified platform to transform how IT and business can turn data into insight. Empower any decision maker with self-service tools to catalog, cleanse, and shape data from any source for use anywhere. Your data experts design the integration rules, while IT governs the use of data and facilitates collaboration across batch, bulk, and master data management scenarios.

## Optimize big data performance in the cloud

Run big data processing when and where you need it—on-premises, hybrid or in the cloud—with the best response time, lowest latency, and most cost-effective use of resources. Build end-to-end big data integration workflows that easily integrate with Amazon Redshift, Elastic MapReduce (EMR), Amazon Kinesis or Azure HDInsight systems, so all your infrastructure runs in the cloud. Or use our powerful self-service tools with Talend Integration Cloud, part of the Talend Data Fabric.

## Protect big data investments with a future-proof architecture

Talend released the first big data integration platform to run MapReduce and Spark on YARN. With each new Hadoop framework, Talend makes it possible to convert ETL and ELT jobs to the latest frameworks with the push of a button, so you can stay ahead of the innovation curve. Subscription pricing, based on users not CPUs or connectors, sets a predictable cost basis even as data volumes and systems grow exponentially.

## Compare

Choose a Big Data Integration solution to best fit your project and budget.



### Download:

#### Free Big Data Sandbox

Go from zero to big data in under 10 minutes

<https://info.talend.com/prodevaltpbdsandbox.html>

Features	Open Studio for Big Data	Big Data	Big Data Platform	Real-Time Big Data Platform
License	Apache	Subscription	Subscription	Subscription
Big Data	Hadoop and NoSQL Components	+ Batch (Spark, MapReduce), Native Hadoop, Connectors	+ Batch (Spark, MapReduce), Native Hadoop, Connectors	+ Real-Time Processing (Spark Streaming), and Machine Learning, High-Speed Messaging, and IoT Connectivity
Design Faster And Scale Easily	900+ Components & Connectors	+ Continuous Delivery, Testing, Sharing, And Debugging	+ Repository Manager	+ Repository Manager
Collaborate Better And Manage More	—	Manage Administration, Deployment, & Automate Tasks	+ High Availability, Load Balancing, And Failover	+ High Availability, Load Balancing, And Failover
Accelerate Data Usage	—	+ Operationalize Data Preparation	+ Operationalize Data Preparation	+ Operationalize Data Preparation
Increase Trust With Data Quality	—	—	Data Profiling, Cleansing, Matching, Masking & Stewardship	Data Profiling, Cleansing, Matching, Masking & Stewardship
Support	Talendforge Community, Help Center Access	+ Guaranteed Response Times, Web And Email Support, Optional 24/7	+ Phone Support, Faster Response, Optional 24/7	+ Phone Support, Faster Response, Optional 24/7
Indemnification/warranty	—	✓	✓	✓

Free Download and Sandbox: [www.talend.com/downloads](http://www.talend.com/downloads)

*“Using Talend, the Otto Group has been able to make quicker and smarter decisions around product lines, improve forecasts, reduce leftover merchandise and importantly, improve our customers’ experience.”*

*Chief BI Platform Architect, Major EMEA Retailer*

## Specifications

Talend Real-Time Big Data supports the following third party components, products and operating systems. Support varies across products.



### Free Whitepaper: Hadoop in the Enterprise

<https://info.talend.com/benchmarkreportheadopenterprise>

#### Supported Operating Systems

- CentOS Linux
- OS X
- Redhat Enterprise Linux
- Solaris
- SUSE Linux
- Ubuntu Linux
- Microsoft Windows

#### Supported Big Data Hadoop Distributions & NoSQL

Amazon Redshift, Amazon DynamoDB, Amazon EMR (including Apache Spark), Apache Hadoop (HBase, HDFS, Hive), Apache Spark, Cassandra, Couchbase, CouchDB, Cloudera Enterprise, Google BigQuery, Hortonworks Data Platform, Microsoft HDInsight, IBM PureData System for Hadoop, MapR, MarkLogic, MongoDB, Neo4J, Pivotal HD, Riak, SAP HANA, Teradata, THD, Vertica

#### Big Data File Format Support

SEQ, JSON, RC, ORC and AVRO

#### Database and Storage Connectivity

Amazon Aurora, Amazon RDS, Amazon Redshift, Amazon S3, AS400, DB2, Derby DB, Exasol, eXist-db, Firebird, Google Storage, Greenplum, H2, HIVE, HSQLDB, Informix, Ingres, InterBase, JavaDB, JDBC, MaxDB, Microsoft OLE-DB, Microsoft SQL Server, MySQL, Netezza, Oracle, ParAccel, PostgreSQL, PostgresPlus, SAP Business Warehouse, SAS, SQLite, Sybase, Teradata, VectorWise, Vertica, Windows Azure Blob Storage

#### Supported SaaS and 3rd Party Applications

Alfresco, Centric CRM, Elasticsearch, JIRA, Marketo, Microsoft CRM and AX, NetSuite, Open Bravo, SAGE X3, Salesforce.com, Salesforce Wave, SAP, ServiceNow, Splunk, SugarCRM, Vtiger CRM

#### Talend Big Data Platform also includes:

- Address Validation, Standardization and Enrichment
- Through a combination of components and services, Talend supports the following address validation partners: Google, Loqate, QAS, Melissa Data and QAS
- Matching Algorithms: Exact Match, SoundEx, SoundEx FR, Levenshtein, Metaphone, Double Metaphone, Jaro, Jaro-Winkler, Q-grams Custom/User-Defined

#### Talend Real-Time Big Data Platform also includes:

- Big Data Supported Messaging Services: Apache Spark Streaming, Apache Kafka, Amazon Kinesis.
- Support for Enterprise Messaging Standards, Transports and other ESB-related Capabilities