



Talend

Industry

Data Management

Website

www.talend.com

Company Overview

Data-driven organizations use Talend's integration solutions to gain instant value from all their data. Natively supporting big data, Talend's open and unified solutions take the complexity out of any integration project and equip IT to be more responsive to the demands of the business, at a predictable cost.

Product overview

Talend Big Data on Cloudera is a powerful and versatile big data integration and data quality solution to extract, load, and process large and diverse data sets for better insights or process optimization.

Solution Highlights

- Easy and fast - Visually develop jobs that run 100% on Spark. A single solution for batch, streaming and real-time data integration and data quality.
- Native big data support - Unlike legacy integration solutions, Talend natively generates Hadoop and Spark code and requires zero-footprint on the Hadoop cluster.
- Connect your legacy IT landscape - Bridges and augments your legacy data assets through 900+ connectors and components for RDBMS, NoSQL, ERP/CRM, SaaS, and all types of flat files.

Cloudera and Talend Deliver Real-Time Big Data

The journey to successful big data management poses many challenges and may create new risks. Acquiring a big data platform, integrating new and existing data sources, managing extreme volumes of data, and ramping up big data skillsets can be very costly and time consuming. Meanwhile, business opportunities may go unnoticed in the volumes of data flowing daily through most organizations.

Now there is a better way to deliver on the promises of big data. The combination of Talend and Cloudera provides a new class of highly scalable, fast big data processing and management across vast amounts of structured and unstructured data, so big data projects can be completed faster, at a lower cost, and using existing skillsets.

Talend's data management environment running in Cloudera Enterprise enables you to create and execute Hadoop integration jobs, process and reconcile big data, and implement data governance processes using an intuitive drag-and-drop interface. This robust environment makes it possible to scale to any level and support any complex data type, so companies can focus on analyzing information instead of manually integrating data.

Big Data Without the Need to Write and Maintain Code

Talend's big data product combines components for Spark, MapReduce 2.0 (YARN), Hadoop, HBase, Hive, HCatalog, Oozie, Sqoop, and Pig into a unified environment so you can quickly load, extract, transform, and process large and diverse data sets from disparate systems.

Talend provides an easy-to-use graphical environment that allows developers to visually map big data sources and targets without the need to learn and write complicated code. Running 100% natively on Hadoop, Talend big data provides massive scalability. Once a big data connection is configured, the underlying code is automatically generated and can be deployed remotely as a zero-footprint job to process the data where it is on your cluster without mandating to move it to an external engine.

Talend provides big data quality functions that take advantage of the massively parallel environment of Hadoop, allowing you to understand the completeness, accuracy, and integrity of data as well as to remove duplicates. Hadoop data profiling allows you to collect information and statistics about big data to assess data quality, repurposing, and metadata. Additional functions of the big data solutions include standardization, parsing, enrichment, matching, survivorship, and monitoring on-going data quality.

Big Data Projects Simplified

Together Talend and Cloudera simplify the task of integrating diverse data structures, no matter the volume or velocity. As the first enterprise data management solution that natively leverages the power and scalability of Cloudera Enterprise, the combined solution provides:

- **High performance** – Talend generates native MapReduce, Hive, Pig and Spark code, provides native integration with Cloudera, leverages Spark Streaming in-memory processing, enables end-to-end Internet of Things interoperability, and is the only data quality solution to run inside of Hadoop. This translates to faster performing applications.

Cloudera Enterprise Benefits

Stores and Analyzes Any Type of Data

- Leverage the full power of your data to achieve pervasive analytics, increase business visibility, and reduce costs
- Bring diverse users and application workloads to a single, unified pool of data on common infrastructure; no data movement required

Enterprise Approach

- Compliance-ready perimeter security, authentication, granular authorization, and data protection through encryption and key management
- Enterprise-grade data auditing, data lineage, and data discovery

Industry-Leading Management and Support

- Best-in-class holistic interface that provides end-to-end system management and zero-downtime rolling upgrades
- Open platform ensures easy integration with existing systems
- Open source to achieve stability, continuous innovation, and portability

Talend Benefits

- Fastest time-to-benefit to the enterprise data hub – Talend dramatically lowers the integration and training efforts needed to bring Hadoop to the core of your IT landscape, accelerating the time-to-value
- Improved data integrity – Native Hadoop data quality helps you discover and cleanse big data and reconcile it with your current processes, business entities, and applications
- Affordable and predictable pricing – Both Talend and Cloudera use an open core and subscription-based pricing model allowing to predictably scale your big data costs
- Future proof – Based on an open source foundation, Talend provides an extensible solution that runs at the innovation pace of big data and Cloudera.

- **Easy, integrated management** – Talend 6.1 is certified on Cloudera Navigator 2.4 (with Cloudera CDH 5.5). This lets users trace data lineage for MapReduce and Spark down to the level of the schema defined by the developer in a data Job, which is crucial for both impact analysis and data lineage.
- **Low technology risk** – Using open standards and the same Eclipse-based development environment, your developers will be familiar with the combined environment. As Talend generates Spark, MapReduce/YARN, Pig Latin and Java, your ramp-up time is shortened.
- **Low financial risk** – Both Talend and Cloudera use a subscription based pricing model, allowing you to start small and scale over time without high up-front financial commitments.
- **Broad connectivity** – Talend provides connectivity to Cloudera from existing and legacy systems with over more than 900 connectors and components for RDBMS, NoSQL, ERP/CRM, SaaS, and all types of flat files.

Turn big data into big benefits with Cloudera and Talend. Not only can Cloudera and Talend support new business initiatives such as real-time analytics, Internet of Things interoperability and insight, customer experience management with sentiment analysis or fraud detection, but can also transform and drive down the costs of your existing IT assets, such as ETL and batch process offloading or Enterprise Data Warehouse optimization.

About Talend

Talend lets organizations unlock all of their data, including historical, live, and emerging data. Through native support of the modern big data platforms, Talend's zero footprint solution takes the complexity out of integration and equips IT to be more responsive to the demands of the business, at a predictable cost.

Based on the most widely used open source integration tools, Talend's enterprise-grade solutions are leveraged by more than 1,300 enterprise customers worldwide. The company has major offices in North America, Europe and Asia, and a global network of technical and services partners. Learn more at talend.com

About Cloudera

Cloudera delivers the modern platform for data management and analytics. The world's leading organizations trust Cloudera to help solve their most challenging business problems with Cloudera Enterprise, the fastest, easiest, and most secure data platform built on Apache Hadoop. Our customers can efficiently capture, store, process, and analyze vast amounts of data, empowering them to use advanced analytics to drive business decisions quickly, flexibly, and at lower cost than has been possible before. To ensure our customers are successful, we offer comprehensive support, training, and professional services.

Learn more at cloudera.com.

cloudera.com

1-888-789-1488 or 1-650-362-0488
Cloudera, Inc. 1001 Page Mill Road, Palo Alto, CA 94304, USA

© 2016 Cloudera, Inc. All rights reserved. Cloudera and the Cloudera logo are trademarks or registered trademarks of Cloudera Inc. in the USA and other countries. All other trademarks are the property of their respective companies. Information is subject to change without notice.