talend*
*open data solutions

Talend Open Studio v3

Talend democratizes data integration.
Through an open and innovative platform, Talend provides a simple, powerful solution.

Proprietary solutions are expensive—many organizations have been forced to address data integration needs by implementing custom programs or routines. More often than not, these integration routines are designed to solve a specific requirement, rather than architected to address a global solution, and without any attempt to leverage already-existing processes. This approach creates maintainability issues, poses robustness challenges, represents a real risk to the integrity and consistency of the information system, and is extremely expensive to maintain.

A versatile data integration solution

Talend Open Studio is a powerful and versatile open source solution that addresses all of an organization's data integration needs:

- Synchronization or replication of databases
- Right-time or batch exchanges of data between systems in the IT infrastructure
- ETL (Extract-Transform-Load) for analytics
- Data migration
- Complex data transformation and loading

Talend Open Studio comprises three major applications (Business Modeler, Job Designer, and Metadata Manager) within a single graphical development environment based on Eclipse, which is easily adapted to corporate needs.

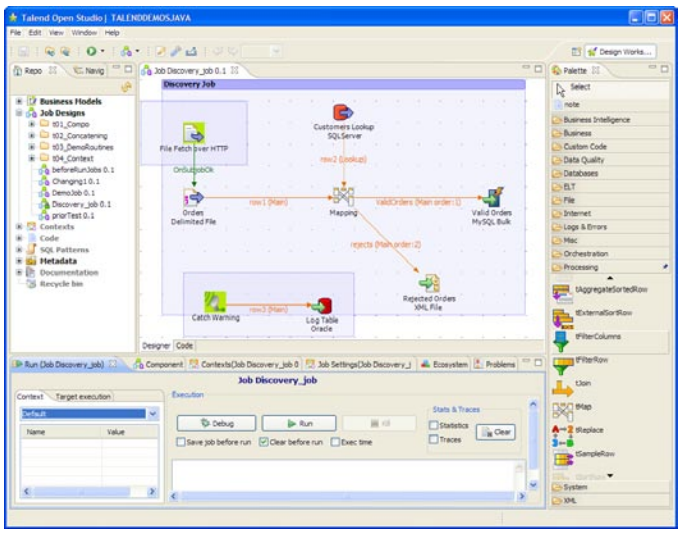
Business-oriented process modeling

Talend Open Studio's Business Modeler leverages a top-down approach, allowing line-of-business stakeholders to get involved in the design of the integration processes and to monitor development progress.

The Business Modeler also regroups all relevant documentation supporting the data integration process in a business-friendly diagram. This a very efficient way of monitoring the jobs and performing impact analysis if a problem arises.

Graphical development

Talend Open Studio's Job Designer provides both a graphical and a functional view of the actual integration processes using a graphical palette of components and connectors— the Component Library.



Integration processes are built by simply dragging and dropping components and connectors onto the workspace, drawing connections and relationships between them, and setting their properties (most properties are inherited from the metadata).

Metadata-driven design and execution

Talend Open Studio is a metadata-driven solution—all metadata is stored and managed in a Metadata Repository, shared by all the modules. The Metadata Repository centralizes all project information and ensures consistency across all integration processes.

Metadata related to source and target systems of the integration processes is easily loaded in the Metadata Repository through advanced database or file introspection, expedited by a number of wizards. The Metadata Repository is based on an open relational model, through which impact analysis can be performed to facilitate maintenance.

Advanced and versatile connectivity

Talend Open Studio offers native technical and business connectors to all IT environments. This wide array of connectors is the key to the successful interoperability of applications and databases; it allows bridging diverse and heterogeneous data structures at unmatched performance rates. It is also continuously expanding, enriching the features of the solution.

Some examples of connectors:
Files: CSV, Excel, positional, XML, LDIF, unstructured
Databases: MySQL, MS SQL Server, DB2, Oracle, Teradata, Ingres, PostgreSQL, EnterpriseDB, Sybase, MS Access, Informix, Firebird
Applications: SAP, SugarCRM, MS Dynamics, Salesforce.com
Log & error: info, warning, statistics, filtering
Data quality: deduplication, fuzzy logic matching, CRC
Transformation: filtering, mapping, reference, aggregation, lookup, XSLT, (de)normalizing
Miscellaneous: Web Services, FTP, HTTP, system, SSH, data generator

See http://www.talendforge.org/components for a complete list of supported connectors.

Talend Open Studio leverages industry-standard languages including Java, Perl and SQL. This allows customers to easily enrich existing components or to create their own, and to plug these components natively in the environment.

Real-time debugging

Talend Open Studio includes powerful testing, debugging, and tuning features that allow the real-time tracking of data flowing through the whole transformation processes, including execution statistics and an advanced trace mode.

And of course, all code generated by Talend Open Studio—regardless of the target language—is always visible and accessible from the design environment.

Deployment and maintenance

Advanced execution context management (test, staging, production, etc.) facilitates integration processes deployment.

Automated documentation generation provides complete and up-to-date technical reference documentation (in XML and HTML).

Robust and scalable execution

Unlike many integration solutions which are based on a centralized integration server, or can only use RDBMS engines to process data, Talend Open Studio distributes the processing across a grid of systems based on their available capacity. These systems do not need to be dedicated to executing integration processes. Instead, Talend Open Studio leverages available resources regardless of their nature.

Talend Open Studio is the only data integration solution that leverages both the traditional ETL (Extract-Transform-Load) approach as well as the ELT (Extract-Load-Transform) approach. ELT leverages the power of the RDBMS engines to execute the data transformations inside the database, achieving unmatched performance for high-volume batches. For each subset of a process, it is possible to choose the most suitable approach, and hence obtain the highest level of performance and scalability.

This architecture design—especially suited to leverage grids (large or small) of inexpensive servers, as well as high-range systems—enables data to be processed at a location closest to its source (thus decreasing data transfers), and maximizes the use rate of computing resources.

> ***Our choice of Talend Open Studio was driven by both technical criteria: the solution is the best fit for our requirements, and by strategic reasons: the Open Source model and Talend's support and services around its solution guarantee our long term investment.***
> **University of Toulouse**

Talend's technology and business vision shatters the traditional proprietary model by providing the flexibility required to meet the data integration and quality needs of all organizations, regardless of their size, level of expertise, or budgetary constraints.

More Information:
www.talend.com
info@talend.com