# Table of Contents

MCF Technology Solutions provides hosted data integration services .............................................. 4  
Creatis Bank manages regulatory compliance with Talend Integration Suite .................................. 6  
Oscar Lambret Cancer Center implements an EMR and analyzes its activities with Talend Integration Suite .................................................................................................................................. 8  
Naville ensures application interoperability ...................................................................................... 10  
The Zannier Group manages its strategic BI project with Talend Integration Suite ......................... 12  
AOL standardizes data integration technology for optimum performance ......................................... 14  
PriceMinister.com automating data exchanges improves reactivity .................................................. 16
Practical Open Source Data Integration
Case Studies & Implementation Examples

Over the past few years, commercial open source vendors have been providing a real alternative to proprietary players. In the data integration space, enterprise grade open source solutions are adopted by many organizations for their data integration and data quality projects. These organizations span all industries, all continents, and all company sizes. More importantly, their projects range from ETL for business intelligence to operational data, data quality, master data management, etc.

This document presents a few selected case studies, illustrating real-life implementations of open source data integration and its associated benefits.
MCF Technology Solutions provides hosted data integration services

The Platform as a Service provider streamlined its integration processes, reduced costs for its clients and enhanced its capability to connect to new data sources with Talend’s data integration technology.

A full service Web 2.0 PaaS service provider

Headquartered in Cleveland, Ohio, MCF Technology Solutions is a full-service Web 2.0 Platform as a Service (PaaS) provider whose role is to help customers maximize cloud computing PaaS tools to create efficient, low cost, and highly effective Business Process Applications (BPA).

MCF has licensing agreements to provide hosted data integration with QuickBase, QuickBooks, Wolf Frameworks, and WorkExpress. Integration jobs are custom designed and executed to process data between clients on the secure Amazon Web Services EC2 service. Currently MCF supports more than 4,500 QuickBase users in a variety of business support applications including Global Supply Chain Management and Sales contact and Project Management. They have been developing QuickBase applications since 2005 and their services include advanced development of JavaScript and .NET tools to extend QuickBase functionality.

Platform as a Service

MCF understood early on that one-size-fits-all software and expensive hardware weren’t options for the smaller companies that comprised their target market. Lower in cost than traditionally developed software tools, and rapidly deployed thanks to PaaS technology, Web-aware applications are emerging as an excellent way for companies to boost the bottom line and build business processes.

“After several years of developing custom integrations with PaaS using traditional technologies,” said Don Larson, Managing Partner of MCF Technology Solutions, “we realized the need to streamline the integration process, reduce costs for our clients and enhance our capability to connect to new data sources.”

MCF originally tried the traditional coding approach (in C#), but it wasn’t efficient and changes were very difficult. The company then looked at Clover ETL, but quickly realized that it would be difficult to self-manage and would make them heavily reliant on Clover’s support organization. Autonomy was key to successfully deploying the solution.

“Fortunately the market for open-source data integration was rapidly emerging, with Talend leading the way,” said Don Larson. “After an exhaustive review of the market, we realized that Talend could provide the support, technology, vision, and partnership to ensure our success in bringing data integration into the MCF

“Talend lets us narrow the gap between the client process need and the technology to support it. The same person having the conversation with our client about their requirements can also build the solution which means no loss in translation, reduced hand-off costs and increased development speeds. The speed at which we can develop and release jobs is significantly better than 4 times faster”

Govind Davis
Solution Designer
offering. The Java-based model gave us confidence in the flexibility and future of the technology, and Talend's base of components and active community means that we can rapidly develop integration processes and quickly get feedback and assistance.”

**Exchanging records economically**

MCF has two business models. For small companies, MCF connects directly to their database. In most cases, the records exchanged aren't large, but make the Talend value proposition available to smaller businesses or departments that would normally not be able to justify the expense. The interface is with typically business process managers who need integration more quickly and at a lower cost than their own IT department can provide. "We might only be pushing 100 records a day," said Govind Davis, Solution Designer at MCF, "but that's extremely valuable for smaller companies. For larger companies, IT involvement is limited to extracting records from the database, and emailing it to MCF, who then extracts the attachments, manipulates the data, inserts it into QuickBase, and sends back the results."

"In the case of large companies," added Govind Davis, "we can offer solutions less expensively than their own IT departments. For example, we're able to charge $30,000 for a solution that would cost $140,000 to implement in-house. It's an important saving, particularly when a small number of data integrations are involved."

"Since we adopted Talend," said Don Larson, "our business has changed dramatically. Previously, we understood that integration was an important part of our solution, but we worried about accuracy and it was very expensive. With Talend, we're able to cut the cost of integration in half. For example, a major area for us has always been integrating between QuickBase and Quickbooks. That used to cost $5,000 and today we're offering it at a base price of $2,500. And the market is huge—there are 400,000 QuickBase users and 4 million Quickbook users."

**Low learning curve, rapid development**

MCF was essentially self-taught on Talend’s data integration technology. "We were able to figure out Talend on our own," said Govind Davis. "In a more traditional development model I would have to meet the customer, write up a specification, and send it to the developer, who would develop the process and send it back to me. Then I'd need to test it. Talend eliminates all these steps. A key benefit for us is that Talend lets us narrow the gap between the client process need and the technology to support it. The same person having the conversation with our client about their requirements can also build the solution which means no loss in translation, reduced hand-off costs and increased development speeds. The speed at which we can develop and release jobs is significantly better than 4 times faster.

**Example of a successful implementation**

Virginia-based ENSAT Corporation provides turn-key environmental services to private industry. MCF provides its service to ENSAT customers as part of a suite of QuickBase applications they deploy for environmental and safety management. One customer, a European Electronics factory, is primarily using the application for managing their factory training schedules and training requirements.

QuickBase provided an ideal solution for the basic system interaction such as scheduling classes; however the requirements analysis requires a complex model of database transformations to determine which employees require which classes based on their inclusion as part of a workgroup as well as their assignment to various tasks associated with their jobs. MCF needed to set up a routine to gather all of the data on tasks, classes, staff and workgroups and output the results of a model that determined compliance requirements and would indicate the status of each staff person in meeting the requirements.

Talend’s data integration technology provided an ideal solution because of its ability to connect and extract data from QuickBase and perform advanced data manipulations (in this case a series of SQL joins). Using Talend, MCF was able to complete the compliance generation routines in less than a week and was able to deploy and monitor the outcomes of the jobs. "This effort would easily have taken 4-6 weeks using code-only technologies," said Govind Davis, "and would not have allowed the direct interaction of our business process experts in the development process which is instrumental to design creativity."
Creatis Bank
manages regulatory compliance with Talend Integration Suite

Managing many databases, the bank depends on Talend’s open source integration solution to feed a Data Warehouse that guides its compliance requirements to the Basel 2 directives.

Specialist in loan consolidation

A specialist in loan consolidation for the past 10 years, Creatis has been part of the Cofidis Group since 2006. Its main activity is the purchase of consumer credit through its retailing channel, via a network of brokers. Creatis has nine agencies distributed over the Nord-Pas-de-Calais and Picardy and provides local banking services and a large range of savings plans to individuals and associations.

Creatis also offers consumer credit via a call center and its website. Today the fast growing company has 250 employees and has developed a multi-channel and multi-product approach to win new customers.

An evolving information system

Creatis needed to reconcile two constraints to further its IT development. In addition to continually evolving the technical architecture, the organization and the processes of the company in coordination with various departments (marketing, risks, sales, financial, etc.), Creatis had to ensure the daily performance of the information system as well as its operational security. Like all banks, Creatis wanted to implement a responsive Information System that would allow it to proactively satisfy its customers.

Therefore, when the bank decided to restructure its mortgage package in 2007, a new information system was developed to manage and market it. The original information system remained in use and evolved to handle other bank products and credit (personal loans, reserve money, etc.) whose migration to the new system would come later. The original database was split into two independent parts. By the following year all bank accounts had migrated to the new Information System based on two software packages—Viveo’s V.bank for banking and Efibank to manage credit—the AS/400 databases and web interfaces developed in Java.

To monitor these activities, Creatis relied on a system developed in-house which was directly connected to the original database. “Because of the new multi-database architecture, data extraction became more complicated and required a lot more work,” explained Stéphane Leclercq, Head of Standards and Development at Creatis. “We quickly became aware that we needed to develop a real data warehouse that would let us efficiently analyze our activities. Our Research and Development department was tasked with identifying appropriate data integration products, focusing on open source solutions.”

We calculate that with Talend Integration Suite we reduced development time by a factor of 3—a particular advantage for teams involved in the migration project. Talend’s solution lets us handle data extractions via a graphical interface that accelerates the production of the documentation and notations that Basel 2 requires.

Stéphane Leclercq
Head of Standards and Development
Talend Integration Suite: a quick response to increased regulatory constraints

Creatis began by testing Talend Open Studio. "In addition to our traditional banking activities, regulatory requirements put in place by Basel 2 required that we reconcile both IT systems and implement a five-year data report. Talend's solutions were particularly effective in this context because of their ability to easily manage multiple databases. After exhaustively testing Talend Open Studio, we opted for Talend Integration Suite whose additional collaborative features allowed our three developers to manage projects as a team."

In addition, the Java-based solution and its Eclipse development interface have accelerated its implementation. "Talend's ease-of-use is a plus," said Stéphane Leclercq, "because there's no downtime spent learning the product. We can focus on business and regulatory aspects instead of wasting time resolving technical issues."

Today, the data warehouse is near completion. It drives all financial activities and bank clients. Each branch manager is now able to get a dashboard of their customers and the central controlling department globally consolidates all financial activities. For example, daily reports indicate contract status and monthly reports help manage various activities, such as litigation.

"Several extractions have been developed to meet our regulatory requirements," said Stéphane Leclercq. "First, we developed a model score with Experian, which helps us calculate profits and losses, in a database dedicated to risk management. We also analyze certain banking activities, such as types of credit cards sold to our customers. Today, most of these extractions are integrated into a spreadsheet, such as Open Office, but we plan to install a free open source viewer—like QlikView—to improve our analytic process."

Development time divided by three

Today, Talend Integration Suite is at the heart of Creatis' monitoring system. "In selecting Talend Integration Suite, we made a strong strategic choice that meshes well with our parent company, Cofidis, which also uses this solution. We didn't want our multi-database architecture to get in the way of data extraction and Talend Integration Suite helped us achieve that," explained Stéphane Leclercq. "We calculate that with Talend Integration Suite we reduced development time by a factor of 3—a particular advantage for teams involved in the migration project. Talend's solution lets us handle data extractions via a graphical interface that accelerates the production of the documentation and notations that Basel 2 requires. Creatis plans to take advantage of Talend's consistency and quality to meet other regulatory requirements. For example, the bank needs to verify data entered by business referrers who directly access its information system.

Creatis takes advantage of Talend's Java libraries to aggregate the data produced by its J2EE application. "We have developed services to produce specific calculations—such as monthly payments for credits—consistent with our internal business rules," concluded Stéphane Leclercq. "In addition, through a customizable system of permissions, our developers can share projects and collaborate more easily than before, increasing productivity. Finally, we are very satisfied with the reliability of the solution and have had no need to contact Talend Support Services. The forums and documentation on Talend's website have enabled us to respond quickly to the few questions we faced!"
Oscar Lambret Cancer Center implements an EMR and analyzes its activities with Talend Integration Suite

The Cancer Research Center consolidates patient medical data within an electronic medical record system and analyzes its activities with the help of a data warehouse.

**A triple mission of treatment, research and education**

Created in 1929, the Oscar Lambret Cancer Center is a private, non-profit, public service hospital. Part of a network of 20 Centers in the Fight Against Cancer within the National Federation of Centers for the Fight Against Cancer (NFCFAC), the Oscar Lambret Center offers its patients an added guarantee of quality and efficiency in its healthcare treatment, education, and research services.

With 750 employees, 250 hospital beds and approximately 16,000 patients (including 3,000 new patients per year), the Oscar Lambret Center annually generates about 37,000 conventional hospital stays, 66,000 outpatient visits, and 49,000 consultations. In the past 5 years, the number of patients treated at the Oscar Lambret Center has increased 11% and the 2006 budget of this publicly-subsidized private establishment, (which is essentially supported by health insurance), rose to nearly 70 million Euros ($102 million).

**An information system comprising a number of business applications**

The Information Management team at Oscar Lambret developed a unique portal for its 110 doctors and scientists to consolidate information from numerous business applications used by its different services (pathology and anatomy laboratories, hematology, nuclear medicine, administrative management, imaging and technical support, etc.) as well as email management tools and meeting management tools. The objective was to establish an electronic medical record (EMR) replicated every 15 minutes, plus an access link to facilitate personnel consultations for the Center’s doctors and healthcare professionals.

The portal was developed by Lille-based Inovelan, a specialist in medical information sharing. Oscar Lambret, in collaboration with Inovelan, tested two data integration solutions—SAP Data Integrator (formerly known as Business Objects), and both versions of Talend’s solution: Talend Open Studio and Talend Integration Suite.

“Both companies offered equivalent functionality, but we chose Talend Open Studio for three reasons,” explained Didier Cauchois, Information Manager at the Oscar Lambret Center. “First, the very user-friendly graphical interface was more intuitive than Data Integrator’s. Second, the Talend solution let us develop in Java, while Data Integrator’s proprietary language would have seriously impacted our learning curve. And third, in calculating the TCO over five years, we realized that there was a clear financial benefit, even if we decided to use Talend’s subscription version, Talend Integration Suite. In addition, Talend is supported by a consistent installed base as well as an extremely active community—very reassuring in terms of product stability.”

**Business Type**
Hospital & Research Center

**Implementation**
Data Consolidation, ETL for Analytics

**Key Facts**
- 70-million Euro budget
- 16,000 patients, including 3,000 new patients per year
- Change Data Capture
- Ease-of-Use (Java-based, graphical interface)
- Enhanced TCO over 5 years
- Data flows developed in less than 5 days
Three Oscar Lambret employees learned enough in a 5-day training session to be comfortable with the tool and develop the first integration flow. Oscar Lambret quickly switched to Talend Integration Suite to replicate transactions (medical and imaging) in its portal from a desktop tool and integrate exterior documents. For example, today, incoming mail, containing test results is automatically scanned and integrated into the portal in PDF format via Talend Integration Suite.

"Talend Integration Suite met our needs perfectly, offering more features and connectors, reactive support services, a unique central repository and Job supervision," explained Franck Craynest, IT Project Leader at the Oscar Lambret Center. "Data is stored in an Oracle database and comes from several business applications: laboratory, medical mail, meetings, traceability, external documents, etc. Today, 90% of our needs are covered; radiotherapy and chemotherapy are being integrated."

**Building a medico-economic data warehouse**

Concurrently, the Center developed a second project to analyze its medico-economic activities. This entailed building a Data Warehouse supported by Axèże software (which runs the data model and allows the generation of cubes) and powered by Talend Integration Suite. The data comes from several business applications, such as hospital stay management, HR, Accounting, etc. The analyses and results are for clinical service managers as well as upper management.

"We want a unique consolidation tool for our integration needs that would capitalize on our knowledge and simplify administration," continued Didier Cauchois. "We also implemented two Talend servers to provide failover capabilities."

The analyses focus on costs and revenue activities, their typology (liver cancer, lung cancer, etc.) and the course of treatment (particularly concerning better organization of patient care). The data warehouse can also be used to produce the statements sent to the Regional Hospital Agency.

**Cost control and functional coverage**

Among the benefits emphasized by the Oscar Lambret Center, the economic aspect is crucial. "We benefit from a rapid ROI and are aware that our projects would have cost a lot more if they had been developed with another tool," acknowledged Didier Cauchois. "Moreover, we’re independent. Our knowledge of the tool meets all our needs and today we can develop a data flow in less than five days, something that was previously impossible. In addition, the tool lets us supervise our data flows."

As part of its integration projects, Oscar Lambret commissioned Talend to develop a specific connector to allow delivery of HL7 format messages—a standard format in the exchange of medical information. The Center plans to give this connector back to the community at no charge.

In the near future, Oscar Lambret also plans to collect data changes in real-time using Talend Integration Suite’s Change Data Capture feature. "In our replication system, we want to automate certain update mechanisms. Change Data Capture lets us track changes to further ensure the reliability of the information that we pass along to our colleagues," concluded Franck Craynest.

"And, contrary to past experiences, we haven’t had any unpleasant surprises using Talend. We have confidence in the success of our future projects."
Naville ensures application interoperability

Talend Integration Suite allows the retail leader to industrialize its data integration.

**The Swiss retail and marketing leader of press and news products**

A subsidiary of Swiss Payot Naville Distribution Group, Naville is owned by Lagardère Services (65%) and Edipresse Group (35%). The company is organized by business, comprising four subsidiaries: Naville Press (distributes 2800 press titles from 23 countries in Switzerland); Naville Détail (manages a chain of 190 shops under the brands Naville, Relay, Press&Web plus a chain of bakery and pastries shops under the Paul brand); Naville Service (supplying local stores with candies, phone cards, etc.); and Naville Livres (distributor—books, travel books and comics). Generating revenues of 300 million Swiss Francs (approx. $260 million), the company employs 300 people in-house and more than 1,000 people in its various shops.

Naville uses heterogeneous information systems. Historically, its four divisions rely on different systems—AS/400; DB2; Internally developed applications for the wholesale and press activities; an Oracle e-Business Suite ERP for the retail activity and accounting; and Microsoft technologies for its B2B websites (B2C soon).

Facing increasing interoperability needs, the company is trying to structure its information systems around two main focuses—an Oracle ERP, and AS/400 for the wholesale activity (the press domain is very specific). "When we decided to migrate our accounting system to the Oracle Financials module of Oracle e-Business Suite, we realized that we needed to develop interfaces to connect that module with other systems downstream and upstream (wholesale logistic application, HR, etc.). We started to develop our new integration structure," explained Dominique Chemin, IT Director at Naville.

At the same time, Naville is dealing more and more often with the need to integrate data flows with numerous partners (banks, press publishers, and various suppliers). The company decided on a single and centralized platform to manage all processes for integration between applications.

**Choosing transparency and real continuity**

After a detailed study of the integration solutions available on the market, Naville tested two proprietary and one open source tool. "We chose Talend Integration Suite based on several strategic criteria. On one hand, cost transparency was critical because we did not want to be chained to a technology without real visibility into its price list. Talend’s prices were much more affordable than those of the proprietary software providers," explained Dominique Chemin. "Talend’s proposal also guaranteed the continuity of the solution with an initial implementation performed by its own teams, followed by knowledge transfer."

---

**Business Type**
Retail

**Implementation**
Data consolidation, operational data integration

**Key Facts**
Application migration
Unique and centralized integration platform
AS/400, DB2, Oracle e-Business Suite ERP, internally developed applications, websites running on Microsoft technologies
Continuity and real cost transparency
Easy to implement and deploy

"**Not only is Talend's performance excellent, but the solution is also easy to implement and deploy; both criteria weighed heavily in the final decision. The basic functionalities met our requirements perfectly. Most of the connectors already existed and the others were developed by Talend in less than a week. That last point let us test the reactivity of the editor and also its listening capacity. Once again, Talend didn’t disappoint us.**"

Dominique Chemin
IT Director
“We also carefully studied the performance of the respective solutions,” continued Dominique Chemin. “Not only is Talend’s performance excellent, but the solution is also easy to implement and deploy—both criteria weighed heavily in the final decision. Finally, the basic functionalities met our requirements perfectly. Most of the connectors already existed and the others were developed by Talend in less than a week. That last point let us test the reactivity of the editor and also its listening capacity. Once again, Talend didn’t disappoint us.”

The IT department of Naville first started its accounting integration project with about 15 application interfaces. “The rapid turn-around on the development of the interfaces is another reason that encouraged us to choose Talend. We cannot afford to waste time because other projects will be implemented as soon as our data integration system is in place,” Dominique Chemin continued.

**A strategic goal—to increase global agility**

That project was strategic for Naville—the increased interoperability will improve its global agility. Other goals are also being pursued. “We want to benefit from improved traceability and better monitoring. Interfacing our monitoring tool (Tango/04) with Talend Integration Suite will make it easier to control the execution and performance of our jobs,” added Dominique Chemin. “Our ambition is also to augment our productivity in terms of interface development and reduce development and maintenance costs. Finally, we want to improve data quality and consistency with Talend which will, in turn, improve the global management of our activities.”

Today Naville has implemented Talend Integration Suite with an initial three years subscription and leverages a Gold Support contract which guarantees high reactivity by the vendor for an unlimited number of incidents. Today, the company manages its integration tasks in batch mode, but it is thinking about using Talend’s solution for its real-time needs as well.
The Zannier Group manages its strategic BI project with Talend Integration Suite

The world leader in children’s fashion enhances its business operational systems with a dashboard to better analyze sales and inventory.

World leader in children’s fashion

Zannier Group is the worldwide leader in children’s fashion with 20 in-house and licensed brands. Founded in 1962 in France, the group demonstrated sustained internal growth and significant acquisitions for more than forty years. Today it has a brand portfolio covering all segments of children’s fashion and is become progressively better known in the adult fashion market.

With a presence in more than 40 countries, Zannier Group generates nearly 50% of its revenue outside of its home country (Europe, the United States, and the Middle East) through more than 800 stores in France and abroad. With world-famous brands and licenses (Absorba, Catimini, Chipie, IKKS, Z, Kenzo, Levis, Marc Jacobs, etc.) Zannier Group posts revenues of more than €650 million ($865 million).

The number of products offered and the compatibility of its collections give Zannier Group a unique market position. Its adaptability to changing fashion trends, its well-researched knowledge of the attitudes and expectations of children and teenagers, and its unique approach to the world of baby clothes and gifts are trademarks that make the Group a leader in the world of children’s fashion. Recently, a number of adult fashion brands were developed which have become major players in this market (IKKS Women, One Step). The reputation of these brands also promotes licensing in various fields—luggage, shoes, watches, perfumes etc.

Guiding activities through analysis of sales and inventory

In 2007, Zannier Group initiated a strategic Business Intelligence project which would add a dashboard to its business operational systems. The plan was to analyze sales and inventory as part of its retail activities (affiliated stores and branch offices) and to gradually expand this functionality to other areas of the company such as mass market retailers and to manage the provisioning of these sectors through Central Purchasing.

The IT Department is developing a data warehouse to consolidate data from the Group’s operational business systems, and particularly from two ERPs designed for the clothing industry—DOTSoft (integrated accounting web software specific to retailing) and Orli, a production management solution used to manage wholesale activities.

“To extract, transform, and load our data into this data warehouse we needed an integration solution powerful enough to handle all our data and flexible enough to integrate seamlessly and transparently with our information system,” explains Didier Dubois, IT Infrastructure Manager. “We thoroughly researched our options and
chose Talend Integration Suite. The solution has many advantages such as ease of operation and implementation, its rich functionality, the ability to integrate specific components and its user-friendly integrated features (automation, reports, logs, incident analysis, etc.)."

**Talend Integration Suite—a major component of Zannier Group's BI chain**

"Talend Integration Suite will be one of the major components of our BI chain. Its strong performance will contribute to a strategic BI project that will accelerate our growth as a company," adds Christian Vincent, Administration/Finance/BI Project Manager. "Beyond its purely technical advantages, we feel that Talend's open source model and the existence of a community participating in product development were important advantages. Finally, the fact that it's Java-based allowed us to make good use of existing skills within the company."

Zannier Group selected Talend Integration Suite in March 2008. After a short ramp up period, development began on the retail activity. Initially, more than twenty data flows were developed and maintained by three in-house developers.

"Although the volumes of data analyzed are not yet comparable to those used daily by mass retailers, Talend Integration Suite offers good scalability that allows us to increase the volumes of data progressively and easily," says Christian Vincent. "Our analyses are structured around four key areas—the consumer, the product, the time, and the store. Each register transaction gives us the information on which we base our reporting. This helps us analyze our sales and stock to improve the services we offer to our end users through our retail network."

**More efficient reporting**

The benefits offered by Talend Integration Suite will be measured through its productivity and ease of use. Although the system is not yet fully implemented, Zannier Group feels that the solution will improve the reporting features provided by the two ERP that it runs.

"Furthermore, thanks to the user-friendliness of Talend Integration Suite, we're letting business users perform queries—currently handled by the IT team. Ultimately this will promote greater autonomy and greater productivity because they won't need to depend on a third-party to analyze new trends," concludes Christian Vincent. "At the same time Talend's solution will help us in the daily operation of our data warehouse. For example, we can run processes automatically every night and the reports from these targeted logs will help us make any necessary adjustments quickly and rerun the jobs with our limited resources."

When production on the first part of the project is finished, the Zannier Group plans to expand the analysis to include its wholesale activities, enabling them to efficiently anticipate market trends and to quickly meet customer needs.
A leader in advertising-supported Web services

AOL is a global ad-supported Web company, with a comprehensive display advertising network in the U.S., a suite of popular Web brands and products, and a leading social media network.

The company’s strategy focuses on growing the size, engagement, and monetization of its worldwide audience by providing highly relevant content and advertising to consumers across the Web. AOL’s businesses include Platform-A, the industry’s largest digital advertising platform; MediaGlow, AOL’s network of Web properties; and People Networks. In addition, AOL continues to run one of the largest Internet access providers in the U.S.

Standardizing technologies

Two years ago, AOL decided that standardizing technology would reduce the need for different integration tools and maximize efficiency among the different development teams. The company began by standardizing their front-end technologies on Apache/Tomcat and their back-end technologies on MySQL. Next, it tackled data integration solutions and ETL tools.

As far as integration is concerned, AOL had been traditionally using a combination of Kettle (Pentaho), eMule, and proprietary tools. They were also using custom developments which created problems related to reuse and consistency.

“We did an evaluation,” says Eric Schvimmer, CTO of AOL’s Commerce and Marketplace Division, “and we knew we needed to meet certain criteria. And, we were sure that we wanted an open source product. As a company we’re moving away from commercial licensing whenever possible. It’s just not a scalable business model.”

Evaluating the options

“Our goals were straightforward,” continues Eric Schvimmer. “It came down to a number of key requirements. First, performance was critical: we needed a fast process to handle small files as well as large files, both simple and complex. Ease-of-use was also important, and we liked Talend’s graphical environment; it’s really a tool that doesn’t need a developer to operate. Extensibility was also important; we wanted to be able to go under the hood and develop additional functionalities if necessary. And finally, robustness—data integration processes are critical to our business and we need a reliable product.”

Talend helps grow the size, engagement, and monetization of AOL’s worldwide audience.

AOL standardizes data integration technology for optimum performance

Of the open source vendors we looked at, Talend was fastest. We found that we can build processes in hours that used to take days to develop. That’s our main benchmark.

Eric Schvimmer
CTO
Commerce & Marketplace
Talend met these requirements and was selected by AOL.

“Another real advantage is that we now have a standard tool that all of our people know how to use,” said Eric Schvimmer. “Currently we have nearly 50 developers using Talend in three locations—India (Bangalore), the United States (Dulles, Virginia) and Ireland (Dublin). By choosing a single tool we can move work around to the most available bandwidth letting people jump from one team to another. It’s part of our standard development toolkit; it’s easy to pick up, and the staff doesn’t need to learn new things.”

Lastly, AOL emphasizes that although Talend is very flexible it stills provides guidelines, framework, and audit capabilities that offer much greater reliability than custom developments.

The advantages of open source

AOL also wanted to standardize on open source because it gave them the ability to make modifications as needed. “We like being part of an open source community,” adds Eric Schvimmer, “because we know the product is going to be updated more frequently than proprietary tools, in part thanks to the contributions from the community. We have forked off a few customizations ourselves that we will contribute back to the community. It’s a system that works really well and we like some of the developments in Talend’s recent 3.0 release.”
PriceMinister.com
automating data exchanges improves reactivity

PriceMinister.com integrates near-real-time price changes and partner inventory—between 1,000 and 2,000 data files per day.

Second-ranked e-commerce site in France

An independent French Group founded in 2000 and employing 180 people, PriceMinister (http://www.priceminister.com) created the Internet concept “Buying and Selling just got safer” in 2001, based on the ‘trusted third-party’ model. PriceMinister puts sellers and buyers in touch providing the necessary guarantees to ensure that all transactions are successful. This means that, the seller is paid and/or the buyer is reimbursed if the sale doesn’t work out.

Today the Group is structured into four major areas:

• new and second-hand products (www.priceminister.com)—20 million visitors per month; 120 million products at reduced prices—PriceMinister is France’s second-ranked e-commerce site
• the automobile industry (www.priceministerauto.com)—300,000 unique visitors, 3.3 million pages viewed
• real estate—(through the ‘A Vendre A Louer’ group, one of the leaders in real-estate ads)
• travel and tourism (www.voyagermoinscher.com)—2 million unique visitors

In 2006, PriceMinister expanded its presence in Europe with the establishment of its Spanish subsidiary, PrimeMinister.es. At the end of 2008, the English version of the website was launched in the UK. Today, the conglomerate claims 9 million members and over 170 million products.

30 million sellers and more than 100 million products

Product volumes range from a few units for individuals, to millions for the largest vendors; and these sellers add or change stock on a daily basis. Because PriceMinister’s 60-person IT team integrates and updates all of this data as it rolls out, automated data feeds were needed to facilitate information processing, and ensure fast online posting.

“Two elements are really critical—how quickly we can import files and how quickly we can set up a new feed,” said Justin Ziegler, Co-founder and Chief Technical Officer at PriceMinister. “We process huge volumes of data and pricing updates are frequent because of PriceMinister’s role as aggregator—a number of sellers offer similar products and that drives the price down. The number of items is also significant—on average, we manage 30 million sellers with more than 100 million items in inventory, and 2 million price changes per day. Typically, we sign up a new partner every day. And between 1,000 and 2,000 data files are updated daily.”

Updating 1,000 to 2,000 data files daily requires reliability and a processing ability that many integration solutions lack, particularly those in this price range. Talend fits seamlessly into our information system, which is critical to our business.

With more than 100 million items on offer, it’s important that we don’t overload our systems and degrade the fast response times that our customers expect.

Justin Ziegler
Co-founder and Information Systems Manager
To manage this data, PriceMinister used software developed in-house which was very expensive to maintain over time. “We thought about developing an exchange format and requiring our sellers to use it; however, this is contradictory to our core strategy which puts no constraints on the sellers,” explained Justin Ziegler. “In addition, since it wasn’t our software, it seemed more efficient to put an established data integration tool in place; that way we wouldn’t have to worry about development and could take advantage of the latest innovations.”

Talend Open Studio—business facilitator

Because the in-house software could neither meet increased demands nor support the XML feeds, PriceMinister tested and implemented Talend Open Studio. “Talend’s solution is a real business facilitator,” said Justin Ziegler. “Previously we had to turn down contracts because the sellers needed to exchange XML files that we couldn’t handle. That’s no longer the case. In addition to XML, we now use Talend Open Studio every time our old system is not up to the task. Importing this data is really critical, because we have more chance to sell if we can offer 20 items instead of 10.”

The ability to quickly absorb any change is also critical for customer satisfaction—prices are frequently lower because PriceMinister favors competition.

Today Talend Open Studio is connected to both an Oracle database containing the 100 million products marketed via PriceMinister and also to FAST, which provides advanced search features. PriceMinister’s IT department is now considering interconnecting the tools to handle data quality issues. “Data quality is clearly a priority for improving our services and we are monitoring and analyzing the quality of all files before integrating them into our database,” added Justin Ziegler. “We now offer customers a report providing details on key indicators—revenue generated and the success rate of the data imports, for example. Clearly, this data quality feature will help improve our interaction with sellers and increase sales.”

Flexible and fast—manages large amounts of data

After using Talend Open Studio intensively for a few months, PriceMinister determined that the product offered three major benefits. “First, the solution lets us integrate data quickly, said Justin Ziegler. “This is critical for several reasons. Customers want maximum reactivity from us to prevent lost sales or to avoid offering an item that is not actually in stock. Outdated information can lose sales and may also result in the loss of both buyers and sellers and thus in lost revenue. We can also add new sellers quickly.”

The second benefit is Talend Open Studio’s ability to support integration operations on huge volumes of data. “Updating 1,000 to 2,000 data files daily requires reliability and a processing ability that many integration solutions lack, particularly those in this price range. Talend fits seamlessly into our information system, which is critical to our business. With more than 100 million items on offer, it’s important that we don’t overload our systems and degrade the fast response times that our customers expect,” said Justin Ziegler.

Finally, Talend Open Studio is highly flexible, reacting quickly to any requirements. “Parameters of new feeds need to be defined in less than a day to satisfy our sellers and to provide fast integration of the updated data. Talend Open Studio gives us the ability to react quickly and to migrate development from one project to another. Our data integration team is currently using the tool and we plan to bring it into general use, in the near future,” added Justin Ziegler. “We have a number of projects coming up—communicating transaction data to our partners and setting up Web Services, just to name two—and Talend Open Studio will play an integral role in our IT plans going forward.”