



Talend Open Studio

Talend democratizza l'integrazione dei dati. Attraverso una piattaforma aperta e innovativa, Talend offre una soluzione semplice e potente.

Molte organizzazioni sono state costrette ad affrontare le esigenze di integrazione dei dati sviluppando programmi o routine personalizzati. Molto spesso, queste routine di integrazione sono progettate per risolvere una specifica necessità, (piuttosto che architettate per affrontare problemi globali) e senza alcun tentativo di migliorare i processi esistenti. Questo approccio pone seri dubbi di affidabilità, rappresenta un rischio reale per l'integrità e la coerenza del sistema di informazione, ed è difficile e costoso da mantenere.

Una soluzione versatile di integrazione dati

Talend Open Studio è una soluzione open source potente e versatile che risponde a tutte le necessità di integrazione dei dati di una organizzazione:

- Sincronizzazione e replicazione di database
- Scambi di dati in tempo reale o differiti
- ETL (Extract Transform Load) per Business Intelligence
- Migrazione di dati
- Complesse trasformazioni e caricamenti di dati
- Qualità dei dati

Talend Open Studio è composto da tre applicazioni principali (Business Modeler, Job Designer e Metadata Manager) in un unico ambiente grafico di sviluppo basato su Eclipse, che è facilmente personalizzabile per specifiche esigenze aziendali. Gli utenti ottengono il supporto dalla Talend Community e integrano i contributi direttamente nel loro ambiente di sviluppo.

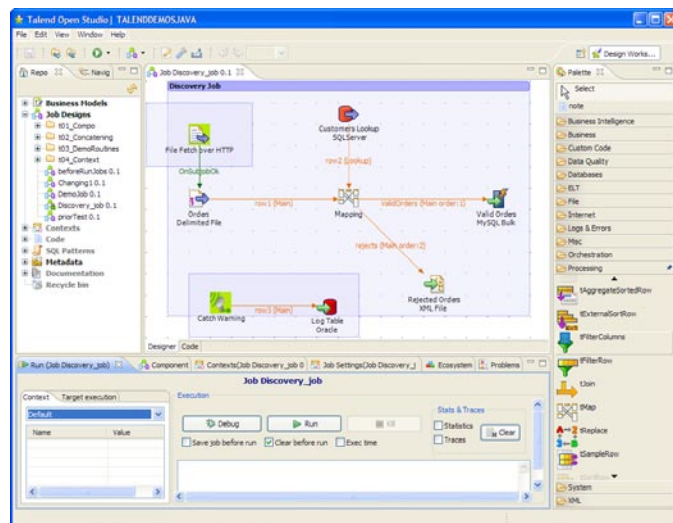
Progettazione dei processi orientata al business

Il Business Modeler di Talend Open Studio utilizza un approccio top-down, consentendo alle parti interessate di essere coinvolte nella progettazione dei processi di integrazione e di monitorare i progressi dello sviluppo. Il Business Modeler raggruppa anche tutta la documentazione rilevante a sostegno del processo di integrazione dei dati in un diagramma business-friendly. Questo è un modo molto efficiente per monitorare i processi di integrazione e per analizzare i potenziali problemi.

Procedimento di sviluppo grafico

Il Job Designer di Talend Open Studio fornisce sia una visione grafica che funzionale dei reali processi di integrazione utilizzando una tavolozza grafica di componenti e connettori. I processi di integrazione sono costruiti con il semplice drag and drop di componenti e connettori sull'area di

lavoro, disegnando le connessioni e le relazioni tra loro, e impostando le loro proprietà.



Componenti e connettori comprendono tutti i tipi di compiti e operazioni sui dati stessi o sulle frequenze di flussi di lavoro. I connettori aiutano l'accesso e la lettura/scrittura dei dati da tutte le fonti a tutti i sistemi di destinazione.

Le proprietà sono configurate centralmente in un'unica schermata al momento della selezione di ogni componente coinvolto nel job o possono essere ereditate dal Metadata Manager.

I componenti complessi sono dotati di interfacce grafiche dedicate e intuitive o da procedure integrate di supporto che aiutano gli utenti a costruire il loro job. Per mantenere la leggibilità del progetto di un job, il diagramma di lavoro può essere diviso in sottoinsiemi, e quindi può essere impostato come sottoinsieme figlio e sottoinsieme padre per stabilirne la sequenza di esecuzione. Altri componenti di orchestrazione, oltre a vari tipi di relazioni aiutano gli utenti a stabilire la sequenza di esecuzione del loro processo. Una console integrata consente agli utenti di monitorare rapidamente l'esecuzione e di rilevare le prestazioni.

Progettazione ed esecuzione basate su Metadata

Talend Open Studio è una soluzione basata su metadati. Tutti i metadati vengono memorizzati e gestiti nel Metadata Manager -il Repository- condiviso da tutti i moduli. Il Repository centralizza tutte le informazioni sul progetto e garantisce la coerenza in tutti i processi di integrazione. Meta-informazioni relative alle fonti e ai sistemi di destinazione dei processi di integrazione sono facilmente

caricate in Metadata Manager attraverso un database avanzato o file di introspezione, facilitato da una serie di procedure guidate. La gestione dei metadati si basa su un modello relazionale aperto, dove le dipendenze dei lavori possono essere facilmente identificate, facilitando il mantenimento della integrazione dei dati, i job di migrazione e la sincronizzazione dei dati. Anche dati contestuali, come i dettagli di connessione al database o percorsi di file possono essere centralizzati nel Metadata Manager, facilitando il loro utilizzo e aggiornamento. Frammenti di codice, procedure o metodi possono anche essere unificati nel repository, facilitando il riutilizzo e il ricalcolo di parti di processo.

Connettività avanzata e versatile

Talend Open Studio offre connettori tecnici e di business nativi a tutti gli ambienti IT. Questa vasta gamma di connettori è essenziale per garantire l'interoperabilità delle applicazioni e dei database, permette di mettere in relazione strutture dati diverse ed eterogenee ad una velocità ineguagliata. E' inoltre in continua espansione, arricchendo le funzionalità della soluzione.

Esempi di connettori:

Files: CSV, Excel, posizionali, XML, LDIF, EBCDIC, multi-schema, destrutturati

Database: MySQL, MS SQL, DB2, Oracle, Ingres, PostgreSQL, Sybase, Vertica, Infobright, ParAccel, Teradata, JDBC, AS/400, eXist, Hive, LDAP, Informix

Applicazioni: SAP, SugarCRM, MS Dynamics, Salesforce, Openbravo, HL7, Talend MDM

Log & error: di informazione, di allarme, statistici e di filtro

Qualità dei dati: deduplicazione, logiche di corrispondenze incoerenti CRC

Trasformazione: Filtraggio, mappatura, referenze, aggregazione, ricercaXSLT, ridefinizione

Internet & Protocolli: Web Services, FTP, HTTP, SSH, POP

Orchestrazione: iterazione, ciclo, replicazione, unificazione

Fare riferimento a <http://www.talendforge.org/components> per la lista completa dei connettori supportati.

Talend Open Studio sfrutta linguaggi standard di mercato che includono Java, Perl e SQL. Ciò consente agli utenti di arricchire facilmente i componenti esistenti o di creare i propri. Una applicazione dedicata per la community - Talend Exchange - consente agli utenti di innestare questi componenti appena creati, nativamente nell'ambiente. Gli utenti possono anche scrivere routine e altri pezzi di codice e memorizzare queste informazioni a livello centrale nel repository per il riutilizzo.

Debugging in tempo reale

Talend Open Studio include potenti funzionalità di testing, debugging e ottimizzazione che consentono il monitoraggio in tempo reale dei dati che fluiscono attraverso gli interi processi di trasformazione, comprese le statistiche di esecuzione e una avanzata modalità di tracciamento.

Naturalmente, tutto il codice generato da Talend Open Studio, indipendentemente dal linguaggio di destinazione, è sempre visibile e accessibile dall'ambiente di progettazione

Distribuzione e manutenzione

La gestione avanzata dell'esecuzione (test, staging, produzione, ecc) facilita la realizzazione dei processi di integrazione. Il caricamento implicito di parametri di contesto direttamente nella progettazione aiuta a sviluppare i diversi ambienti di esecuzione e gestirli facilmente. La distribuzione dei processi attraverso i sistemi aziendali può essere facilmente effettuata sotto forma di servizi dati o i servizi di integrazione dati, tramite il comodo strumento di esportazione. La generazione automatizzata di documentazione fornisce una completa e aggiornata documentazione tecnica di riferimento (in XML e HTML), che aiuta diversi utenti e persone coinvolte a mantenere e aggiornare i processi ereditati. La funzione di rilevamento delle dipendenze aiuta gli utenti ad identificare le dipendenze tra i processi di integrazione sviluppati in Talend Open Studio e semplifica l'aggiornamento globale di un gran numero di processi immagazzinati centralmente nel Repository.

Esecuzione robusta e scalabile

A differenza di molte soluzioni di integrazione che sono basate su un server di integrazione centralizzato, o che possono utilizzare solo i motori RDBMS per processare i dati, Talend Open Studio consente agli utenti di esportare i processi in file eseguibili che possono essere distribuiti attraverso una rete di sistemi o esposti come Web Services. Questi sistemi non hanno bisogno di essere dedicati per l'esecuzione di processi di integrazione. Invece, Talend Open Studio sfrutta le risorse disponibili, indipendentemente dalla loro natura.

Talend Open Studio fa leva sia sul tradizionale approccio ETL (Extract-Transform-Load) che sull'approccio ELT (Extract-Load-Transform). ELT sfrutta la potenza dei motori RDBMS per eseguire la trasformazione di dati all'interno del database, ottenendo prestazioni senza eguali per la elaborazione sequenziale di grandi volumi di dati. Per ogni sottoinsieme di un processo, è possibile scegliere l'approccio più idoneo, e quindi ottenere il massimo livello di prestazioni e scalabilità.

Questa architettura, che è particolarmente adatta a sfruttare sia reti di server poco costosi, che di sistemi di gamma alta, permette l'elaborazione dei dati in una locazione più vicina alle fonti (diminuendo così i trasferimenti di dati), e massimizza il tasso di utilizzo delle risorse di calcolo.

Una attiva Comunità di utilizzatori

Gli utenti di Talend Open Studio possono contare su una vasta e attiva comunità di utenti attraverso gli strumenti di community online TalendForge:

- Forum: <http://www.talendforge.org/forum/> è il luogo dove tutte le domande possono trovare una risposta.
- Wiki: <http://www.talendforge.org/wiki/> fornisce l'installazione e le informazioni generali.
- Bugtracker: <http://www.talendforge.org/bugs/> è lo strumento per segnalare qualsiasi problema con Talend Open Studio.
- Exchange: <http://talendforge.org/exchange/> è qui che i contributi esterni possono essere condivisi.
- Babili: <http://www.talendforge.org/babili/> è il portale di localizzazione in cui tutti i prodotti Talend possono essere tradotti.

Grazie ai contributi comunitari, Talend Open Studio è ora disponibile in 15 lingue.

Informazioni: www.talend.com | info@talend.com