

## Data Integration Features Comparison Matrix

Talend Open Studio

vs

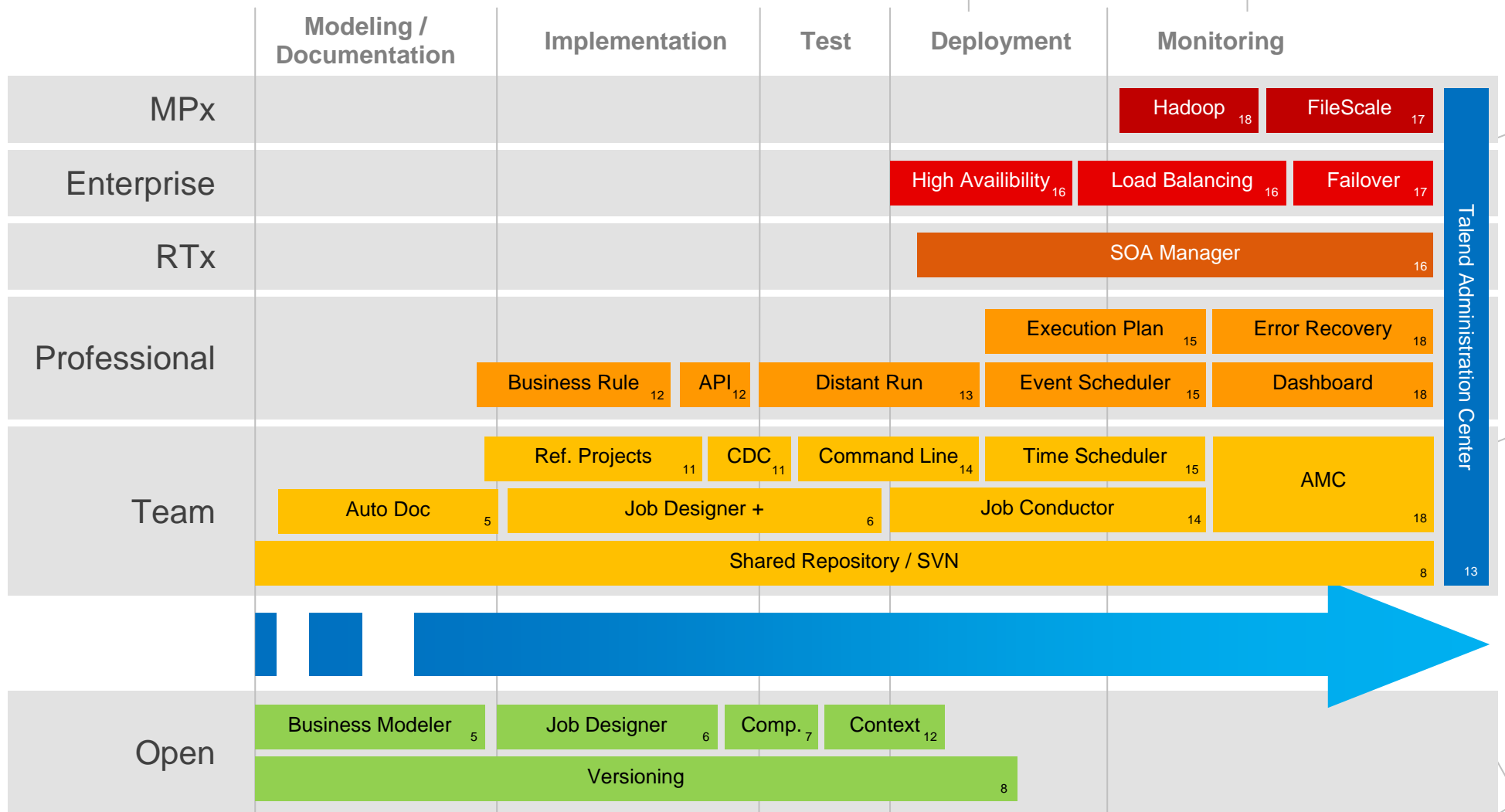
Talend Integration Suite

Team, Professional, RTx, Enterprise & MPx editions

<http://www.talend.com/products-data-integration/matrix.php>

Q1 2011

# Feature details



## Tipo di licenza e indennizzi

### Licenza GPL Open Source

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx

La GNU General Public License è una licenza che stabilisce le condizioni legali per la distribuzione di software libero di un progetto GNU.

Lo scopo della licenza GNU GPL è quello di garantire i seguenti diritti per l'utente:

- Il diritto di eseguire il software per qualsiasi utilizzo e senza limitazioni;
- Il diritto di analizzare il funzionamento del software e adattarlo alle proprie esigenze.

Se l'autore delle modifiche del software decide di distribuire questo software, deve farlo sotto la licenza GPL.

L'intero testo della licenza GPL può essere consultato all'indirizzo: <http://www.opensource.org/licenses/gpl-2.0.php>

### Accesso al codice sorgente

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx

Il codice sorgente di Talend Open Studio è disponibile a: <http://www.talendforge.org/trac/tos/>

Per la completa trasparenza e coerenza, Talend fornisce inoltre ai propri clienti che richiedono il codice sorgente con tutti gli strumenti disponibili nella versione commerciale.

### Licenza a sottoscrizione

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx

Le versioni "enterprise" includono caratteristiche a valore aggiunto (vedi sotto) e Servizi che migliorano i prodotti open source, queste versioni sono distribuite con una licenza commerciale.

Il modello di pricing di Talend garantisce la trasparenza e la prevedibilità: il prezzo non è basato sui volumi di dati o su eventuali ulteriori esigenze di connettori o CPU, ma è basato sul numero di sviluppatori, il livello delle funzionalità (versione selezionata) e termini di sottoscrizione.

Questo approccio ad abbonamento garantisce il ritorno sugli investimenti: il numero di licenze può essere aumentato o diminuito ogni anno per adeguarlo alla evoluzione di un progetto e al numero di sviluppatori impegnati.

Le soluzioni Talend sono più economiche da implementare, mantenere e supportare, sono dal 50 all' 80% meno costose rispetto alle soluzioni proprietarie equivalenti.

## Indenizzo

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx

Poiché i software open source sono il risultato dello sforzo di sviluppo collaborativo, il codice finale unisce i contributi di diverse risorse (sviluppatori). Se l'integrazione dei vari contributi al codice non è attentamente gestita e controllata, l'uso del software finale potrebbe violare il copyright di software proprietari.

L'utente finale potrebbe quindi essere soggetto ad azioni legali e finanziarie per violazione, anche se tale violazione non è stata intenzionale.

Talend offre una clausola di indennizzo ai suoi clienti alle versioni commerciali dei propri prodotti. Questa è una garanzia per l'utente che Talend fornirà una protezione giuridica e finanziaria, nel caso in cui il codice di Talend violasse i diritti di terzi.

## Supporto e Documentazione

### Comunità Talend: Forums, Bugtracker...

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx

La comunità degli utilizzatori di Talend è composta da decine di migliaia di professionisti, ed è estremamente attiva. I principali contributi della comunità includono:

- I test e la qualità delle nuove versioni,
- le richieste di nuove funzionalità,
- la traduzione e localizzazione dei prodotti,
- il sostegno e gli scambi attraverso il forum,
- lo sviluppo e la condivisione di nuovi componenti, connettori, jobs, modelli e altri plug-in....

Talend Exchange consente ai membri della comunità di pubblicare i propri plug-in al fine di condividerli con altri utilizzatori. La maggior parte di questi contributi sono in ultima analisi, integrati nel prodotto, dopo gli approfondimenti e i miglioramenti apportati dal nostro team interno di R&D.

Inoltre, Talend contribuisce a molti fondamentali progetti open source ed è un membro delle Fondazioni Eclipse e Apache.

### Supporto di livello aziendale con SLA

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx

Con la sottoscrizione di Talend Support Services, potete beneficiare dell'esperienza dei nostri esperti tecnici, che sono quotidianamente in contatto con il nostro team R & D. Questi servizi sono stati istituiti per assicurare efficacia, sicurezza e tranquillità ai clienti sottoscrittori dei nostri prodotti commerciali. Sono disponibili in due livelli: Gold e Platinum. A ciascuno di questi livelli è associato un tempo di risposta garantito in relazione a un bug dichiarato, ed un tempo di risposta necessario per fornire una patch, ecc.

<http://www.talend.com/professional-support/support.php>

## User Guide, Reference Guide

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx

La documentazione di Talend Open Studio è disponibile come download gratuito in formato PDF, in inglese e francese. Due guide, la Guida per l'utente (276 pagine) e le componenti Guida di riferimento sono disponibili all'indirizzo: <http://www.talend.com/resources/documentation.php>

È inoltre possibile acquistare una versione stampata di quelle guide su Amazon:

[http://www.amazon.com/Talend-Open-Studio-Guide-Version/dp/1453632077/ref=sr\\_1\\_1?ie=UTF8&s=books&qid=1296040035&sr=1-1](http://www.amazon.com/Talend-Open-Studio-Guide-Version/dp/1453632077/ref=sr_1_1?ie=UTF8&s=books&qid=1296040035&sr=1-1)

Inoltre, è possibile dare un'occhiata al nostro tutorial che è una fonte preziosa di informazioni!

E' possibile accedervi tramite: <http://www.talendforge.org/tutorials/menu.php>

## Modeling / Documentazione

### Business Modeler

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx



The Il Business Modeler è uno strumento non tecnico (come Microsoft Visio). Ti aiuta a strutturare tutta la documentazione pertinente e gli elementi tecnici a sostegno del processo di integrazione dei dati in un diagramma business-friendly consentendo ai diversi team (design, Dev, Test, Prod. ...) di lavorare su un modello comune, utilizzando uno strumento comune.

Ad esempio, gli utenti aziendali utilizzano i business models per esprimere le loro esigenze di integrazione dati. Gli sviluppatori e lo staff operativo IT possono così meglio comprendere queste esigenze di business e tradurli in processi tecnici (Jobs). Al termine di ogni fase di realizzazione tecnica (Jobs), il business models può essere facilmente aggiornato, mostrando i progressi di sviluppo per gli altri soggetti interessati.

Gli amministratori di database (DBA) possono utilizzare il business models per condividere i necessari metadata per la connessione DB e il system architect può così avere il quadro d'insieme delle esigenze richieste in termini di integrazione dei dati.

Progettare il business models fa parte delle best practices che le organizzazioni dovrebbero adottare in una fase iniziale dei progetti di data management o integrazione, al fine di garantirne il successo. Poiché i business models in genere contribuiscono a rilevare e risolvere rapidamente i colli di bottiglia e i punti deboli del progetto aiutando a limitare le spese superflue e/o ridurre l'investimento iniziale.

### Auto Doc

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx

Questa funzionalità consente di generare, su richiesta, una documentazione tecnica dettagliata per tutti i job. Questa documentazione raccoglie i metadata dei job (autore, versione, stato, data di aggiornamento, ecc), una visualizzazione grafica del job e tutti i parametri di tutti i componenti utilizzati in questo job in un formato interattivo di facile utilizzo (HTML / XML).

Questa documentazione può essere facilmente arricchita con commenti personalizzati.

## Auto Doc+

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx



Con autodoc +, la documentazione tecnica (vedi paragrafo precedente) viene generata automaticamente per ogni versione di ogni job: quando si salva un lavoro, la sua documentazione viene aggiornata e memorizzata nel repository, quindi, viene automaticamente condiviso e reso disponibile a tutti gli utilizzatori.

Autodoc + permette inoltre la personalizzazione della visualizzazione grafica di questa documentazione, aggiungendo il proprio logo e il nome della vostra azienda, o cambiando i colori attraverso un foglio di stile CSS.

## Implementazione

## Job Designer

*this feature is included in:*

Open

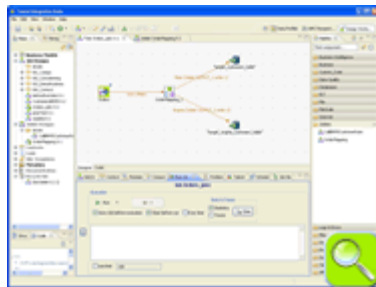
Team

Professional

RTx

Enterprise

MPx



Il Job Designer fornisce i processi di integrazione sia tramite una visione grafica che funzionale utilizzando una tavolozza grafica dei componenti e dei connettori.

I processi di integrazione sono costruiti semplicemente trascinando i componenti e i connettori su uno spazio lavorativo grafico, creando connessioni e relazioni tra loro e impostando le loro proprietà.

Tramite una libreria completa di componenti, le caratteristiche del Job Designer consentono l'accesso a tutti i tipi di sorgenti e destinazioni necessari per i processi di integrazione, migrazione o sincronizzazione dei dati.

I componenti e i connettori includono tutti i tipi di operazioni sui dati stessi, sulla gestione dei dati, nonché sul sequenziamento del flusso di dati. I connettori sono di aiuto per l'accesso e la lettura / scrittura di tutti i dati e i sistemi di destinazione per l'integrazione, la migrazioni e la sincronizzazione dei dati. I parametri sono configurati centralmente in una sola schermata per

la selezione di ogni componente coinvolta nel processo o possono essere ereditati dal Metadata Manager (Repository).

I componenti complessi sono dotati di interfacce grafiche dedicate e intuitive o procedure guidate integrate per aiutare gli utenti a costruire il loro job.

Per mantenere la leggibilità del progetto di un job, il diagramma di lavoro può essere diviso in Subjobs, e quindi possono essere definite come job figlio e job padre per la sequenza della loro esecuzione. I componenti di orchestrazione, nonché vari tipi di relazioni aiutano gli operatori a fissare la sequenza di esecuzione dei processi. Una vista della console integrata consente agli utenti di controllare rapidamente l'esecuzione, controllare e monitorare le prestazioni direttamente dallo Studio.

## Componenti

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx



Talend offre connettori tecnici e di business open source nativi, per accedere a tutti gli ambienti IT. Questa vasta gamma di connettori è la chiave per il successo dell'interoperabilità di applicazioni e database; consente di connettere strutture dati diverse ed eterogenee ad una velocità di prestazioni senza pari. E' inoltre in continua espansione, arricchendo le soluzioni open source di integrazione, migrazione e sincronizzazione di Talend.

Più di 550 componenti sono disponibili gratuitamente, il 60% dei quali sono progettati e sviluppati dalla comunità Talend.

I connettori e i componenti sviluppati esternamente possono essere condivisi attraverso Talend Exchange (<http://talendforge.org/exchange/>). Un certo numero di componenti presentati passano attraverso la validazione e l'ottimizzazione di Talend, prima che siano nativamente integrati e supportati.

Fate riferimento a <http://www.talendforge.org/components> per un elenco completo dei connettori supportati.

## Supporto ETL

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx

ETL (Extract, Transform & Load) è la modalità di default utilizzata dalle soluzioni di integrazione dati di Talend. Essa consiste nel processare righe di dati, una dopo l'altra, in un modello di flusso. Questa modalità è specificamente adattata per essere utilizzata in ambienti eterogenei e consente l'integrazione di dati presenti in qualsiasi formato dai sistemi di origine a quelli di destinazione (servizio web, file, database, MOM, applicazioni aziendali, ecc.).

La modalità di ETL può essere utilizzata sia in batch che in tempo reale. I processi ETL possono essere eseguiti in parallelo per accelerare ulteriormente la loro esecuzione.

L'architettura unica di Talend non si limita a qualsiasi motore di esecuzione in quanto genera processi autonomi che possono essere implementati su qualsiasi server (interno o esterno all'azienda). Inoltre, i processi ETL possono essere eseguiti il più vicino possibile ai dati minimizzando il tempo di accesso e riducendo al minimo il consumo di banda, oltre a eliminare i colli di bottiglia.

Nello stesso Job, questo approccio può essere combinato con l'approccio ELT (vedi paragrafo successivo) per ottenere il massimo livello di prestazioni senza vincoli architetturali.

## Supporto ELT

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx

Le soluzioni di data integration di Talend supportano anche la modalità ELT (Extract, Load & Transform), che consiste nella processare i dati in una serie di operazioni definite (utilizzando le funzioni di Union, Except e Intersect) direttamente sul DBMS del database di destinazione.

Questa modalità è riservata per l'utilizzo in un ambiente omogeneo (un database). Ha il vantaggio di sfruttare le risorse materiali disponibili ed è particolarmente indicata per il trattamento di grandi volumi di dati in "Data Warehouse Appliance" in ambienti come Teradata, Netezza, ecc...

Nello stesso Job, questo approccio può essere combinato con l'approccio ETL (vedi paragrafo precedente) per ottenere il massimo livello di prestazioni senza vincoli architetturali.

## Versioning

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx



Le versioni degli elementi nel Talend Studio possono essere facilmente gestite attraverso l'uso delle funzionalità manuali native di versioning.

Un numero di versione maggiore e minore è impostato automaticamente al momento della creazione del job, e può quindi essere facilmente incrementato nel corso del tempo e degli aggiornamenti, tramite l'apposito pannello di controllo di versione disponibile direttamente nella Designer perspective di Talend Studio.

Tutti gli elementi creati in Studio possono quindi essere distinti da un numero di versione: Modelli di Business, Jobs, routine, metadati e documentazione ...

Il versioning è generalmente parte di un regime di best practice che quando necessario, mira ad agevolare la riusabilità nonché il ritorno ad una fase di sviluppo precedente.

## Shared Repository

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx



La Shared Repository (o Metadata Manager) è progettata per consolidare tutte le informazioni e i metadati aziendali del progetto in un repository centralizzato condiviso da tutte le parti interessate ai processi di integrazione.

Su Studio, gli utenti hanno accesso ai progetti secondo i loro ruoli e le autorizzazioni definite in Talend Administration Center.

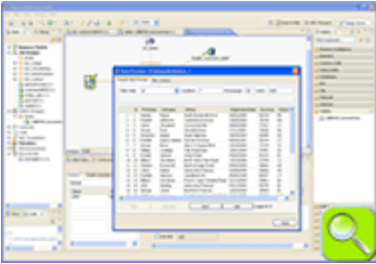
Questo archivio condiviso permette così il lavoro di squadra e la collaborazione tra tutte le persone coinvolte in un progetto di integrazione. Aiuta a conservare e condividere tutti i loro oggetti Talend: Modelli di Business, Jobs (processi), Joblets, routine, definizioni di metadati (come ad esempio la connessione a sistemi sorgente/destinazione) ...

Dietro il repository condiviso risiede un sistema di standard industriale per la gestione delle sorgenti (Subversion) che consente l'archiviazione e la gestione di tutte le versioni di tutti gli elementi.

Un sistema di bloccaggio automatico garantisce che il job che è stato progettato è di fatto bloccato e che nessun altro utente potrebbe cambiare lo stesso job nello stesso momento.

Dalla versione 4.0, abbiamo sfruttato appieno la potenza del versioning , per consentire di gestire diversi rami, check-in/check-out, commit manuale o automatico, commenti ...

|                    |                                     |      |      |              |     |            |     |
|--------------------|-------------------------------------|------|------|--------------|-----|------------|-----|
| <b>Data Viewer</b> | <i>this feature is included in:</i> | Open | Team | Professional | RTx | Enterprise | MPx |
|                    |                                     |      | ●    | ●            | ●   | ●          | ●   |



Mentre sviluppate dei Job con Talend potreste aver bisogno di vedere il contenuto di diversi sistemi sorgente o destinazione (file, DB, ecc). Il DataViewer vi aiuta ad entrare in profondità nei sistemi sorgente e di destinazione indipendentemente dalla applicazione normalmente necessaria per accedervi: Notepad per file testo e csv query browser SQL per tabelle di database, MS Excel per i files .XLS, html browser, etc...

Non vi è alcuna necessità di disporre di diversi strumenti o di sistemi browser per accedere in profondità nei sistemi laddove sono presenti i dati. Il Data Viewer utilizza il percorso definito sorgente/destinazione per accedere direttamente ai dati.

Il Data Viewer può farvi risparmiare un sacco di tempo dato che questo strumento è accessibile direttamente all'interno dello Studio, attraverso un semplice click del tasto destro del mouse su qualsiasi componente. E 'un modo pratico per visualizzare i dati contenuti nei vostri sistemi sorgente/destinazione, indipendentemente dal loro formato (Excel, tabella DB, CSV ...) mentre si stanno sviluppando i vostri processi di integrazione.

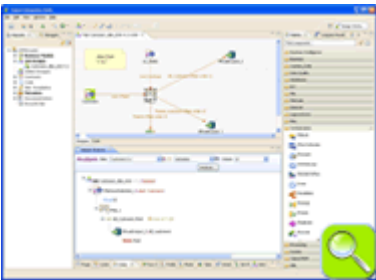
|                |                                     |      |      |              |     |            |     |
|----------------|-------------------------------------|------|------|--------------|-----|------------|-----|
| <b>Wizards</b> | <i>this feature is included in:</i> | Open | Team | Professional | RTx | Enterprise | MPx |
|                |                                     |      | ●    | ●            | ●   | ●          | ●   |

|                       |                                     |      |      |              |     |            |     |
|-----------------------|-------------------------------------|------|------|--------------|-----|------------|-----|
| <b>Dynamic Schema</b> | <i>this feature is included in:</i> | Open | Team | Professional | RTx | Enterprise | MPx |
|                       |                                     |      | ●    | ●            | ●   | ●          | ●   |

Gli schemi dinamici consentono la progettazione di jobs con numero e strutture di colonne sconosciuti. A seconda della scelta dello sviluppatore, le colonne dinamiche possono essere mappate direttamente alla struttura di destinazione utilizzando la modalità Pass-through.

L'applicazione principale di tale funzionalità può essere uno scenario di replica o un semplice mappatura uno-a-uno delle diverse colonne. Questa caratteristica rende facile la progettazione di questi tipi di jobs. Ad esempio, uno sviluppatore che ha bisogno di migrare un intero database con centinaia di tabelle può farlo senza conoscere tutte le strutture delle tabelle utilizzando un unico job!

|                        |                                     |      |      |              |     |            |     |
|------------------------|-------------------------------------|------|------|--------------|-----|------------|-----|
| <b>Impact Analysis</b> | <i>this feature is included in:</i> | Open | Team | Professional | RTx | Enterprise | MPx |
|                        |                                     |      | ●    | ●            | ●   | ●          | ●   |



La funzione Impact Analysis aiuta a capire quali potrebbero essere le conseguenze di un cambiamento. Questa funzione è disponibile nel Metadata Manager. È possibile eseguire un Impact Analysis su qualunque colonna di qualunque metadata (database, file ...). In un rapporto grafico e interattivo, il risultato della Impact Analysis mostra dove si può rintracciare la colonna e vedere tutte le operazioni ad esso applicata, dalla sorgente alla destinazione attraverso il job.

È possibile esportare il report come file HTML.

## Data Lineage

*this feature is included in:*

Open

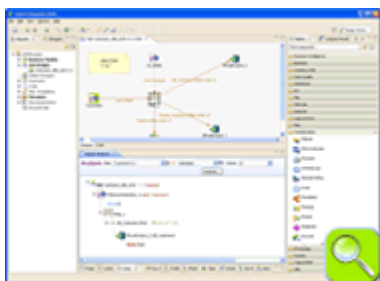
Team

Professional

RTx

Enterprise

MPx



La funzione di Data Lineage ti aiuta a capire dove si è verificato un cambiamento.

Questa funzione è disponibile nel Metadata Manager e può essere effettuata su qualunque colonna di metadati (DB, file). Il risultato del data lineage viene mostrato in un rapporto che traccia un cambiamento, a ritroso dal sistema di destinazione a quello di sorgente di un job.

È possibile esportare il report come file HTML.

## Job Compare

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx



La funzione Job Compare aiuta a identificare le differenze tra due versioni di un job o di jobs differenti.

Job Compare è completamente integrato in Talend Integration Suite Studio. Il risultato di Job Compare è un rapporto visivo e interattivo in HTML o XML in cui sono evidenziate le differenze.

In questo esempio, il rapporto del confronto mostra che il campo di delimitazione delle proprietà del componente tFileInputDelimited non è definito nello stesso modo per entrambi i jobs comparati: nella versione 3.2 il delimitatore è "\ t", mentre nella versione 4.2 è "\ n".

## Joblets

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx



I Joblets aiutano a fattorizzare una parte di job (o Subjob) in un componente Joblet. Basta selezionare i componenti che formano la parte di lavoro che avete bisogno di riutilizzare o si vuole fattorizzare e cliccare sulla voce di menu "refactoring per Joblet".

Automaticamente, il job design viene semplificato, nel momento in cui i componenti selezionati vengono compressi in un unico componente Joblet. Questo componente Joblet può essere condiviso attraverso la cartella dedicata Joblets nella palette dei componenti ed è quindi facilmente riutilizzabile in qualsiasi altro job.

I Joblets semplificano drasticamente il mantenimento dei jobs ridondanti e complessi.

Inoltre, il meccanismo di "Impact Analysis" ti aiuta a trovare i jobs che utilizzano uno specifico Joblet.

## Reference projects

*this feature is included in:*

Open

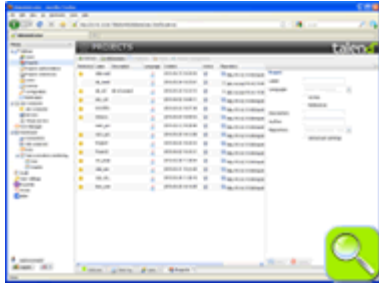
Team

Professional

RTx

Enterprise

MPx



I Reference Projects contribuiscono ad evitare la duplicazione (copia-incolla) di oggetti (Jobs, Routine, Documentazione, Metadata ...) tra i progetti.

I progetti "Slave" sono legati a uno (o più) progetto (i) "Master" e quindi ne ereditano gli elementi/oggetti.

Le risorse provenienti dal progetto Master appaiono nel progetto Slave in modalità di sola lettura: sono disponibili per il riutilizzo e l'esecuzione, ma non possono essere modificate.

Poiché un forte legame è così stabilito tra progetti Slave e Master, quindi, non appena qualcuno modifica un elemento del progetto di Master, tutti i progetti Slave vengono aggiornati di conseguenza.

La Reference Projects condivide tutti gli elementi ridondanti di un progetto (job, modelli, metadati), al fine di renderli disponibili ad altri progetti. Questa funzionalità consente di sfruttare e riutilizzare il 30% di elementi che di solito sono comuni a tutti i progetti di integrazione dei dati, riducendo drasticamente la manutenzione associata.

## Change Data Capture

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx



Il data warehousing implica l'estrazione e il trasferimento dei dati da uno o più database in uno o più sistemi di destinazione per l'analisi. Tuttavia, questo significa l'estrazione e il trasferimento di enormi volumi di dati che potrebbero richiedere un elevato dispendio, di risorse e di tempo.

La funzionalità di catturare solo i dati modificati in tempo reale, è conosciuta come Change Data Capture (CDC). Catturare le modifiche riduce il traffico di dati tra sistemi e contribuisce a ridurre i tempi di ETL.

L'architettura CDC di Talend è basata su un modello di pubblicazione/sottoscrizione (publisher/subscriber). Il publisher acquisisce le modifiche di dati e le rende disponibili ai sottoscrittori (Talend Jobs). I sottoscrittori utilizzano le modifiche apportate ai dati ottenuti da parte del publisher.

Questa funzione rileva i record modificati in tempo reale, permettendo ai dati modificati di essere inviati immediatamente al Jobs sottoscrittore, riducendo di conseguenza il tempo necessario a caricare e aggiornare i dati durante l'ETL o la data integration operativa.

Talend Change Data Capture realizza le modalità più comunemente utilizzate: Trigger e Redo log. La modalità a disposizione dipende dal tipo di database interessati.

## Business Rules

*this feature is included in:*

Open

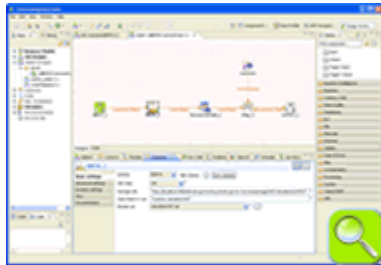
Team

Professional

RTx

Enterprise

MPx



I Business rules sono generalmente definiti dagli utenti business attraverso i documenti delle specifiche che vengono poi interpretati e realizzati da personale tecnico.

Talend Integration Suite incorpora un motore di regole business che aiuta gli utenti a configurare le proprie regole di business. Gli utenti possono quindi definire i criteri di segmentazione del mercato (per età, regione ...) e impostare le loro regole di business attraverso un foglio di calcolo Excel o attraverso l'interfaccia Drools Guvnor direttamente nel web-based Talend Administration Center.

L'interfaccia Drools Guvnor permette di utilizzare agli esperti di business, un editor grafico per creare e modificare le regole in modo rapido e diretto, controllare l'accesso a norme e le altre caratteristiche, gestire le versioni delle regole e modificarle nel tempo. Le regole possono essere testate e richiamate dai jobs sviluppati. (vedi slide successiva)

## API

*this feature is included in:*

Open

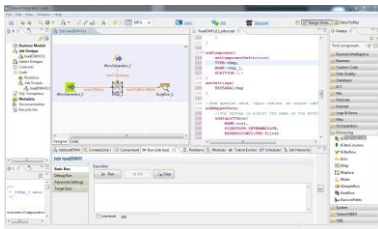
Team

Professional

RTx

Enterprise

MPx



Sophisticated but simple to grasp scripting language which enables programmatic job generation

Fully featured dedicated editor including syntax highlight, code completion aid & warning system

Automatic reverse engineering from an existing Job design to produce the corresponding script

## Test

### Context Management

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx



I context consentono di esternalizzare quasi tutti i parametri di componenti/jobs . Questo, ad esempio, aiuta gli utenti per definire i parametri al volo, in fase di esecuzione o di utilizzare impostazioni diverse per i test e per la produzione.

I Context possono essere definiti come necessari per tutti i tipi di ambienti (sviluppo, test, produzione ...) senza alcuna limitazione in termini di numero di context creati.

Gli utenti possono muoversi da un context all'altro in qualsiasi momento, in fase di progettazione o di esecuzione per utilizzare l'impostazione definita.

I valori dei parametri possono essere modificati tramite una finestra di dialogo in fase di progettazione e collaudo. Inoltre, una componente dedicata al caricamento dei parametri può essere utilizzato per sovrascrivere dinamicamente qualsiasi valore.

## Distant Run

this feature is included in:

Open

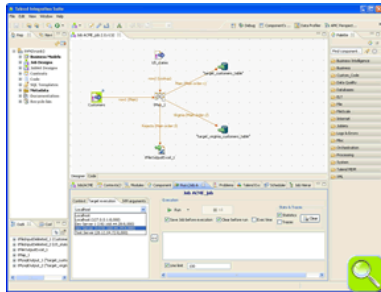
Team

Professional

RTx

Enterprise

MPx



La funzione Distant Run permette l'esecuzione remota di jobs su qualsiasi server direttamente dallo Studio.

Questo può essere estremamente utile quando è necessario testare i jobs, ad esempio:

- in una configurazione simile a quella di un ambiente di produzione,
- su vari sistemi operativi,
- a richiesta, su sistemi specifici

in quanto evita di passare attraverso procedure di implementazione complesse.

Il sistema di destinazione può essere selezionato in modo dinamico in fase di esecuzione direttamente da Studio. In questa modalità di esecuzione remota restano disponibili tutti i debug regolari e le opzioni di statistiche tracce e real time.

## Deployment

### Talend Administration Center

this feature is included in:

Open

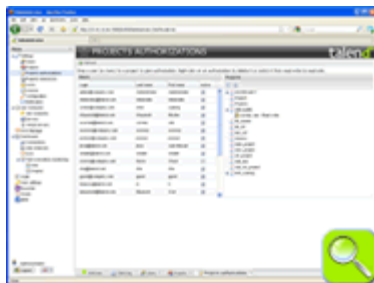
Team

Professional

RTx

Enterprise

MPx



Tutte le offerte di prodotti Talend a sottoscrizione contengono uno Studio (o un numero maggiore a seconda del numero degli utilizzatori) e una parte di software che può essere installata su un server e amministrata attraverso una interfaccia web-based, il **Talend Administration Center**.

Tutti gli Studios sono quindi collegati ai progetti definiti nel Talend Administration Center non più in locale ma in remoto.

Talend Administration Center è un'applicazione leggera (in un browser, senza necessità di implementazione), che aiuta i PM di progetti di integrazione per amministrare gli utenti, i progetti, i privilegi, le utenze, le licenze ...

L'autorizzazioni del progetto sono assegnate facilmente utente per utente (supportando le directory LDAP). Agli utenti sono quindi concessi i diritti per accedere ai progetti in base al loro ruolo: Nessuna autorizzazione, sola lettura, lettura e scrittura ...

Gli utenti possono poi condividere gli oggetti dei repository (Jobs, modelli di business, i metadati connessione DB ...) con altri utilizzatori, direttamente nel loro Studio, per i progetti per i quali sono autorizzati. Maggiori informazioni sul repository condiviso nelle diapositive seguenti.

A seconda del Talend Integration Suite Edition cui si è iscritti, numerosi plug-in aggiuntivi sono disponibili sul pannello di navigazione posto a sinistra (Dashboard, manager SOA, manager Server ...).

## Job Conductor

*this feature is included in:*

Open

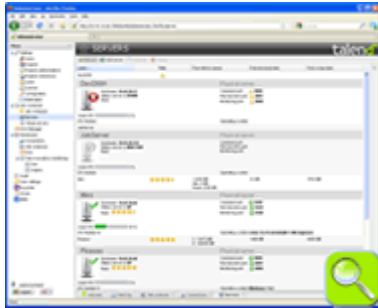
Team

Professional

RTx

Enterprise

MPx



Il Job conductor coordina l'esecuzione dei lavori di integrazione dei dati. Fornisce un'interfaccia centralizzata di esecuzione da cui tutti i job possono essere avviati su richiesta in seguito a schedulazioni basate sul tempo (da Team Edition) o sugli eventi (dalla Professional Edition).

Il modulo Job Conductor si basa sul "JobServers " o agenti, che sono piccole applicazioni che vengono installati su ogni server su cui saranno eseguiti i jobs.

Dopo che gli agenti sono configurati, il Job Conductor permette di monitorare, in tempo reale, tutte le risorse hardware (CPU disponibile, RAM, HD ...) aiutando a distribuire le esecuzioni di jobs attraverso una griglia, sulla base del server maggiormente disponibile. Il supporto nativo JMX consente di monitorare più di 40 indicatori. Qualunque job può quindi essere implementato su qualsiasi server mediante un solo click!

## Command Line

*this feature is included in:*

Open

Team

Professional

RTx

Enterprise

MPx



I processi di integrazione sviluppati con il Job Designer possono essere implementati, aggiornati ed eseguiti al di fuori della GUI Talend Studio, utilizzando il modulo **Command Line**.

Il modulo **Command Line** di Talend fornisce una serie di opzioni da linea di comando che permette agli sviluppatori e agli amministratori di eseguire facilmente operazioni batch.

Quasi tutte le funzioni di gestione di job offerti attraverso il Talend Studio e il Talend Administration Center sono disponibili anche attraverso la Command Line. Ciò comprende, ad esempio funzioni quali: l'aggiornamento delle proprietà dei Job, promuovendo la migrazione di progetti in produzione, l'esportazione/importazione di jobs o gruppi di jobs, ecc

La funzione Command Line rende facile e veloce l'implementazione di numerosi e complessi job includendo le loro dipendenze e metadata di esecuzione.

L'Help nativo del command line fornisce un elenco esaustivo di tutti i comandi disponibili con una breve descrizione della funzione.

## Time Scheduler

this feature is included in:

Open

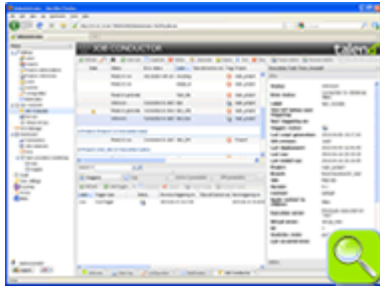
Team

Professional

RTx

Enterprise

MPx



Il Time-based scheduler ti aiuta ad avviare l'esecuzione di un job in un orario/date definite (primo Lunedì del mese, ogni Martedì ...) o su base regolare, per un periodo di tempo. Un Task è utilizzato per centralizzare tutte le informazioni necessarie per l'esecuzione del job (nome del progetto, nome del job, versione del job, server ...)

Il task è quindi attivato secondo la schedulazione e il job è così implementato ed eseguito automaticamente sul server definito al momento definito. Un comodo sistema di controllo dello stato aiuta a monitorare lo stato di attivazione e il successo di esecuzione o di fallimento del job direttamente dal job conductor.

Nella versione Professional Edition è disponibile un ulteriore schedulatore basato su eventi/files. (vedi la slide Event scheduler).

## Event Scheduler

this feature is included in:

Open

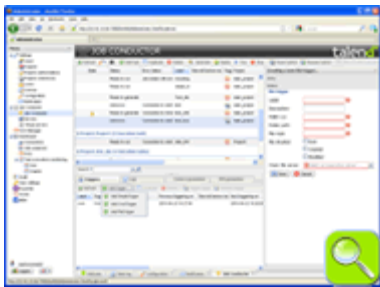
Team

Professional

RTx

Enterprise

MPx



L' **Event Scheduler** estende le capacità di programmazione basata sul tempo per l'integrazione in tempo reale.

L'ascoltatore di eventi (Event listener) consente ai processi di esecuzione di avviare una esecuzione su richiesta o in base ad eventi.

Gli eventi possono essere basati su file come ad esempio all'apparire, allo scomparire o alla modifica di un file oppure basato su SQL utilizzando la condizione "wait for". Una volta che l'evento atteso è identificato, l'attività di esecuzione viene attivata e l'implementazione del job viene effettuata.

È possibile aggiungere facilmente nuovi avviamenti (trigger) basati su eventi ad ogni Task estendendo la industrializzazione delle esecuzioni automatiche.

## Execution Plan

this feature is included in:

Open

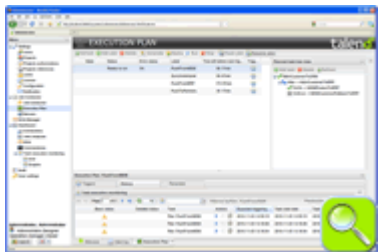
Team

Professional

RTx

Enterprise

MPx



La funzione Execution Plan aiuta a sequenziare e orchestrare le esecuzioni dei vari job e facilitare il recupero dell'errore, direttamente dal Job Conductor. L' Execution Plan è una funzionalità basata su tasks che descrive le dipendenze tra i diversi task orchestrando la sequenza di esecuzione.

Le dipendenze delle attività sono definite attraverso una visione gerarchica dei compiti principali e secondari in cui ogni attività può avere un ruolo subordinato.

Gli Execution Plan possono essere schedulati, avviati e possono utilizzare parametri di esecuzione definiti dall'ambiente da questa singola vista del Job conductor.

## SOA Manager / Webservices

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



Il SOA Manager offre un ambiente web-based completamente grafico per esporre uno o più Jobs di integrazione dei dati sotto forma di servizi (Web Services), consentendo la loro implementazione automatica all'interno e tra applicazioni e sistemi eterogenei utilizzando binding SOAP (RPC o basati su documenti). Una procedura WSDL guidata aiuta a generare descrizioni WSDL per esporre lavori come Web Services e trovare le corrispondenti UDDI entries quando vengono utilizzati i Web Services.

La funzionalità SOA Manager offre capacità avanzate di gestione della richiesta in ingresso sulla base di un sistema di code e di insiemi ottimizzati. L'insieme di active services definito dall'utente gestisce una serie di richieste in tempo reale, mentre il gestore di code gestisce le richieste aggiuntive, regolando il flusso per una elaborazione asincrona.

Il SOA Manager aiuta a esporre i job progettati in Talend Integration Suite Studio in un ambiente di servizio Web-centrico e stabilisce in tal modo l'interoperabilità tra più applicazioni e sistemi.

## Load Balancing

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



Il modulo di **Grid Conductor** (accessibile tramite il Job conductor) ottimizza la scalabilità e la disponibilità dei processi di integrazione, assicurando un utilizzo ottimale della griglia di esecuzione.

Il Grid Conductor si basa sulla definizione di server virtuali, che raggruppano le risorse isponibili, indipendentemente dal tipo di sistema (CPU, OS ...).

I task sono assegnati ai server virtuale del Grid Conductor, piuttosto che ad ogni singolo server di esecuzione.

Attraverso un monitoraggio costante delle risorse disponibili sui server di esecuzione, il Grid conductor garantisce che tutti i jobs siano eseguiti facilmente al tempo prefissato e sfruttando le risorse disponibili, eliminando i colli di bottiglia creati dal tradizionale approccio basato su singoli server.

Questo rimuove le preoccupazioni relative alla prelazione della risorsa quando un gran numero di jobs vengono eseguiti simultaneamente, o quando vengono utilizzati server non dedicati. Grid Conductor fornisce inoltre un fail-over automatico nel caso in cui una risorsa di esecuzione non sia più disponibile.

## High Availability

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



L' High Availability è ottenuta attraverso la capacità di utilizzare più Job Conductors e server per l'esecuzione del job.

D'altra parte, la clusterizzazione dei database garantisce il failover e previene qualsiasi interruzione nell'esecuzione.

|                 |                                     |      |      |              |     |              |       |
|-----------------|-------------------------------------|------|------|--------------|-----|--------------|-------|
| <b>Failover</b> | <i>this feature is included in:</i> | Open | Team | Professional | RTx | Enterprise ● | MPx ● |
|-----------------|-------------------------------------|------|------|--------------|-----|--------------|-------|

|                  |                                     |      |      |              |     |            |       |
|------------------|-------------------------------------|------|------|--------------|-----|------------|-------|
| <b>FileScale</b> | <i>this feature is included in:</i> | Open | Team | Professional | RTx | Enterprise | MPx ● |
|------------------|-------------------------------------|------|------|--------------|-----|------------|-------|

**Benchmark Results**

**Configuration**

- 2 CPU Xeon 5520 Quad Core / 2.26GHz / 24 GB RAM
- Sun StorageTek 2540 Fiber Channel: 12 HD SAS 136Gb

**Results**

|              | 1         | 10           | 100           | 300         | 550         |
|--------------|-----------|--------------|---------------|-------------|-------------|
| # rows       | 6 million | 60 million   | 600 million   | 1.8 billion | 3.3 billion |
| data volume  | 740 MB    | 7.4 GB       | 74 GB         | 229 GB      | 415 GB      |
| sorting time | 6 sec     | 1 min 02 sec | 21 min 30 sec | 2 hr 01 min | 5 hr 17 min |

Independent benchmark made by Sun Micro

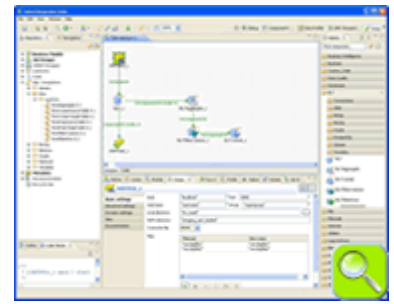
Talend Integration Suite MPx è ideale per quelle architetture multi-server, multi-CPU, e multi-core in cui il codice ed i distinti sotto-processi possono essere eseguiti in parallelo per sfruttare al meglio l'architettura. Questa funzionalità di gestione massiva e parallela, ottimizza la capacità dei server aziendali e il numero di processori disponibili, migliorando notevolmente il tempo di elaborazione.

Talend Integration Suite MPx offre la tecnologia unica **FileScale** che sfrutta l'architettura hardware del server di esecuzione e massimizza le prestazioni di algoritmi di ordinamento di basso livello. La tecnologia FileScale lavora in modo (molto) massiccio su file di grandi dimensioni. Sfrutta al massimo l'architettura di esecuzione in quanto non è limitata dalle restrizioni di JVM o dalle limitazioni dei motori tipici delle architetture tradizionali di data integration.

La tecnologia FileScale ordina e trasforma i dati utilizzando algoritmi matematici innovativi ad alte prestazioni per l'elaborazione dei dati. Esso sfrutta l'architettura MapReduce per scomporre automaticamente qualsiasi operazione di trattamento dei dati in una serie di processi granulari, ottenendo grandi prestazioni. Vedi Sun Microsystems workbench:

[http://blogs.sun.com/aja/entry/talend\\_s\\_new\\_data\\_processing](http://blogs.sun.com/aja/entry/talend_s_new_data_processing)

|               |                                     |      |      |              |     |            |       |
|---------------|-------------------------------------|------|------|--------------|-----|------------|-------|
| <b>Hadoop</b> | <i>this feature is included in:</i> | Open | Team | Professional | RTx | Enterprise | MPx ● |
|---------------|-------------------------------------|------|------|--------------|-----|------------|-------|



Apache Hadoop è un framework software Java open source che supporta le applicazioni ad alta densità di dati distribuite. Sfrutta l'architettura Map Reduce e permette alle applicazioni di lavorare con migliaia di nodi e petabyte di dati utilizzando grandi griglie di server economici. Talend Integration Suite MPx include un supporto nativo per Hadoop permettendo di scalare a qualsiasi livello e di supportare qualsiasi tipo di dati complessi, per cui le aziende possono sfruttare la loro cluster Hadoop per i picchi di volumi di dati e per le trasformazioni complesse.

Un set dedicato di componenti disponibili nella tavolozza help leggono e scrivono sia sistemi HDFS che sistemi Hive e includono template di funzionalità ELT e SQL.

# Monitoring

## AMC

this feature is included in:

Open

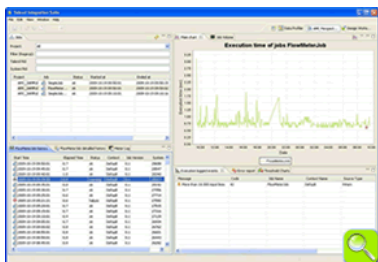
Team

Professional

RTx

Enterprise

MPx



**Talend Activity Monitoring Console** è una comoda interfaccia grafica ed uno strumento di supervisione centralizzato.

Fornisce dettagliate capacità di monitoraggio che possono essere utilizzate per consolidare i log di informazioni raccolti, comprendere la interazione tra i job sottostanti, prevenire guasti che potrebbero essere generati inaspettatamente e supporta le decisioni di sistemi di management .

Activity Monitoring Console monitorizza I job events (successi, fallimenti, allarmi,etc.) I tempi di esecuzione e I volume di dati attraverso una singola console da un punto centralizzato.

Questo strumento è disponibile in formato stand-alone o può essere pienamente integrato nello Studio.

## Dashboard

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



Il Dashboard è una versione basata su Web della Activity Monitoring Console che può essere acceduta facilmente attraverso un Web browser.

La Dashboard fornisce diagrammi sulle prestazioni di esecuzione e indicatori di stato, che consente ad ogni utilizzatore autorizzato di visualizzare sia lo stato attuale che quello storico della esecuzione di qualunque processo di integrazione.

Fornisce inoltre dettagliate capacità di monitoraggio che possono essere utilizzate per consolidare i log di informazioni raccolti, comprendere la interazione tra i job e i componenti sottostanti, fornisce informazioni tempestive sulla esecuzione dei Task, previene guasti che potrebbero essere generati inaspettatamente e supporta le decisioni di sistemi di management.

## Error Recovery

this feature is included in:

Open

Team

Professional

RTx

Enterprise

MPx



I processi di esecuzione dei jobs possono richiedere molto tempo, così come le operazioni di backup e ripristino. Talend Integration Suite Studio include una funzionalità di checkpoint di ripristino che è impostato in fase di progettazione di lavoro.

In caso di guasto, i processi possono essere ripresi da uno dei checkpoint. Gli sviluppatori possono anche progettare e integrare specifiche gestione di errori in risposta a specifiche condizioni di errore utilizzando le istruzioni della funzione "on-failure" del checkpoint.

I checkpoints di recupero possono essere adeguatamente avviati a intervalli specifici del flusso di dati (alle connessioni di avviamento). Lo scopo è di ridurre al minimo la quantità di tempo e lo sforzo necessario quando un processo di esecuzione deve essere riavviato a causa di un fallimento.

Con l'aiuto della funzionalità di error recovery checkpoint, il processo può essere riavviato dall'ultimo checkpoint prima del fallimento (o da qualsiasi altro checkpoint prima che il guasto si verificasse), piuttosto che dall'inizio del processo di esecuzione del lavoro.