

LA DATA INTEGRATION SECONDO **SAS**

Il processo di Data Integration è molto importante e spesso determina il successo di un progetto di Business Intelligence. SAS propone SAS Enterprise Data Integration Server, in grado di rispondere a tutte le esigenze per l'integrazione e qualità dei dati. Può accedere a tutti i tipi di dati, offre potenti routine per la Data Quality ed è integrato con gli altri strumenti di SAS9 Enterprise Intelligence Platform

di **Michele Lovati**



Enzo Ferrari, responsabile tecnologie di Data Integration di SAS

La maggior parte delle aziende dispone di un sistema informativo basato su piattaforme eterogenee e applicativi che gestiscono formati di dati diversi: gli applicativi aziendali, spesso, sono stati installati con esigenze diverse, con una visione limitata alla realtà dipartimentale o locale, spesso non condivisa centralmente. In più, di solito sono stati sviluppati in tempi diversi, con scelte di strumenti e piattaforme influenzate dalle diverse ondate tecnologiche. Negli anni si sono susseguite molteplici stratificazioni di piattaforme, applicazioni, tool e, più in generale, di tecnologie informatiche cui occorre dare forma, riducendone la complessità.

Per sopravvivere alle turbolenze del mercato di oggi, in azienda è sempre più necessario avere una vista unica e integrata, sincronizzata e in tempo reale su tutti i dati eterogenei (operazionali, analitici e transazionali): per avere informazioni utili a disposizione dei decisori, quindi, le diverse fonti di dati devono essere integrate, e spesso la necessaria pulizia di dati, che garantisce precisione, coerenza e completezza, richiede notevoli sforzi per analizzare le informazioni archiviate nei singoli sistemi, e cogliere le interrelazioni tra le diverse componenti.

La Data Integration è una fase molto delicata, che determina spesso il successo delle soluzioni implementate a valle, solitamente sistemi di Business Intelligence: in una recente indagine condotta su un campione di aziende europee, il 66% degli intervistati ha risposto che la scarsa pulizia dei dati incide in modo rilevante sulla redditività di un'impresa. È chiaro che dati di buona qualità assicurano alle imprese un vantaggio competitivo e una base solida sulla quale impostare le decisioni e i processi di business.

La fase di Data Integration non consiste soltanto nel mettere insieme archivi aziendali, come ci spiega **Enzo Ferrari, responsabile tecnologie di Data Integration di SAS** (www.sas.com/italy): «La fase di Data Integration non permette soltanto di collegare dei dati tra loro, ma anche di interoperare con i sistemi esistenti, predisponendo processi costanti per l'integrazione, che consentono l'accesso ai dati in tempi e modi appropriati, assicurando un elevato livello di qualità alle informazioni target: in sintesi, la Data Integration è in grado di trasformare i dati in informazioni». SAS Enterprise Data Integration Server, parte integrante di SAS9 Enterprise Intelligence Platform, è stato aggiornato da

poche settimane ed è progettato per consolidare in modo rapido ed efficiente i dati aziendali provenienti da molteplici fonti, permette di gestire processi Etl di caricamento e aggiornamento, e integra routine avanzate sia per la valutazione dei dati sia per la loro bonifica e qualità. Il Server, inoltre, è in grado di gestire i metadati dall'inizio alla fine del flusso dati e nel passaggio attraverso le diverse strutture informative, integrandoli sia con gli strumenti di Data Quality, sia con le componenti analitiche e di fruizione: questa integrazione con le altre componenti di SAS9, oltre a rafforzare una gestione "collaborativa", è un punto di forza della soluzione, e una caratteristica distintiva della società, che da trent'anni si occupa di analisi dei dati e di Business Intelligence con un'offerta completa, e unica, per tutta la catena dell'informazione. «La nostra è una soluzione che soddisfa appieno le richieste di chi si occupa di Data Integration – aggiunge Ferrari (enzo_ferrari@ita.sas.com) –, garantisce che i dati raccolti e inviati siano completi, gestisce i processi di validazione e arricchimento dei dati, assicura eccellenti livelli di servizio sia per quanto riguarda le tempistiche di consegna e la messa in esercizio delle soluzioni strategiche, permettendo la documentazione di tutti i processi di Data Integration, in modo da semplificare il riutilizzo delle informazioni e la manutenzione, con un forte risparmio dei costi. In più, è in grado di evolvere velocemente, sia nelle dimensioni sia nelle capacità operative». SAS Enterprise Data Integration Server accede direttamente, senza dover scrivere alcuna riga di codice ad hoc, a tutti i tipi di dati, contenuti in qualsiasi file, Db, Rdbmx e Erp, compresi i dati legacy, su una vasta gamma di piattaforme e di sistemi operativi: su questi dati permette una visualizzazione standardizzata delle differenti strutture, e un sistema visuale di selezione del flusso dei dati. È inoltre in grado di archiviare le informazioni in qualunque formato: Db, Rdbms, SAS Storage. Un'altra caratteristica distintiva di SAS riguarda la fase di Data Quality, per la quale offre strumenti unici e integrati: profiling interattivo, Data Parsing, Data Standardization e Data Matching, che permettono di analizzare i dati di partenza, trovando anche eventuali difetti (per esempio valori inattesi, nulli o fuori range), di generare automaticamente le regole per identificare i record da correggere, formattare, normalizzare o eliminare, in caso di duplicazione, evitando che dati errati possano essere caricati nel sistema di destinazione. Il supporto delle regole di normalizzazione nazionali, la ricchezza delle funzioni predefinite, la possibilità di personalizzare le regole da applicare nel processo di trasformazione, le funzionalità predittive e i tool di applicazione dei modelli di data mining ne fanno uno strumento unico, riconosciuto come la miglior soluzione da numerosi analisti indipendenti. «SAS Enterprise Data Integration Server – chiarisce Ferrari – è la tecnologia più innovativa per rendere il processo di Data Integration maggiormente efficace ed efficiente, riducendo sensibilmente i tempi necessari alla sua realizzazione e rilevando, al contempo, eventuali problemi e ostacoli nella fase di disegno piuttosto che in quella di sviluppo. La nostra soluzione fa in modo che i clienti si preoccupino meno dell'accuratezza, della disponibilità e dell'affidabilità dei loro dati affinché possano concentrare maggiormente i loro sforzi sull'utilizzo dei dati: questo consen-

te di raggiungere i migliori risultati sia nella Data Integration sia nella Business Intelligence, fin dalle fasi iniziali del processo». La potenza dello strumento non è solo legata alla sua ricchezza funzionale.

Come tutte le soluzioni SAS, il server ha un'architettura aperta, è scalabile a livello enterprise, è multiplatforma e multi-thread. In più, ha il supporto automatico per ambienti di grid e distributed computing, grazie al quale le aziende possono ottimizzare l'impiego delle risorse hardware e software disponibili e gestire volumi di dati maggiori in tempi più brevi, senza la necessità di installazioni hardware supplementari.

Inoltre, SAS Data Integration Server consente una capacità di analisi del cambiamento senza precedenti, grazie a funzionalità di data lineage e impact analysis che possono tenere traccia dei cambiamenti intervenuti nei dati.

«Gli addetti alla Data Integration – spiega Ferrari – sono ora in grado di individuare tempestivamente anche il più piccolo cambiamento intervenuto in un modello dati e la sua possibile influenza su un intero progetto. In questo modo aumenta notevolmente l'efficienza del processo, perché vengono risparmiate ore di lavoro nel processo di trasformazione, aggregazione, arricchimento e instradamento dei dati, dall'inizio al termine del flusso di integrazione, sia in modalità batch sia in tempo reale». La produttività è aumentata anche grazie a SAS Data Integration Studio, la console di comando full client per le fasi di Data Integration, con specifici wizard per guidare le attività di accesso ai dati, creare le tabelle, i job, i flussi, le regole di trasformazione. SAS Data Integration Server è orientato allo sviluppo in team e al riutilizzo, e offre strumenti validi per il mantenimento e l'organizzazione dell'intero processo, avanzati tool di debugging, un miglior supporto collaborativo, che contribuiscono a ridurre i tempi di sviluppo e ad abbattere i costi di manutenzione. I processi di trasformazione dati sono integrabili anche con strumenti di schedulazione esterni, e possono consentire la Data Integration in tempo reale, alla base del real-time warehousing. Le avanzate funzionalità offerte da SAS Enterprise Data Integration Server permettono inoltre l'invocazione di servizi di integrazione,

«SAS Enterprise Data Integration Server è la tecnologia più innovativa per rendere il processo di Data Integration maggiormente efficace ed efficiente»

trasformazione e arricchimento dati in Real Time, favorendo la realizzazione di una efficace Architettura Orientata ai Servizi (Soa). «Nella nuova versione – conclude Ferrari – SAS ha fatto di tutto per ottimizzare l'impiego dell'infrastruttura It, massimizzare la produttività complessiva e ridurre i costi d'esercizio. Abbiamo rilasciato uno strumento sicuramente avanzato, e focalizzato alla soluzione di complesse problematiche di business, in grado di lavorare al massimo anche con volumi di dati crescenti e ridotte finestre di elaborazione degli stessi, per rilasciare progetti di Data Integration nel pieno rispetto delle tempistiche e dei livelli di budget previsti».