

GESTIONE INTEGRATA DEI DATI

Nuove necessità organizzative impongono di gestire i dati in modo più efficiente e di semplificare i processi e l'infrastruttura IT

Feedback: redazione-cbr@edizionibig.it

DI GIANCARLO MAGNAGHI

Le grandi e medie organizzazioni gestiscono i propri processi in modo sempre più articolato, con il conseguente proliferare di applicazioni con logiche e dati differenti. Molte applicazioni non sono in grado di comunicare tra loro e costituiscono delle isole di automazione, dette anche silos di informazione o stovepipe (tubi di stufa). Questa mancanza di comunicazione produce inefficienze, poiché gli stessi dati vengono memorizzati in luoghi diversi ed è impossibile automatizzare i processi trasversali rispetto ai silos di informazione, ma necessità organizzative impellenti come l'aumento della velocità di reazione e l'agilità nel modificare i processi e i modelli di business, impongono di gestire i dati in modo più efficiente e di semplificare i processi e l'infrastruttura IT per raggiungere un livello superiore di trasparenza, di controllo e di affidabilità.

I problemi di integrazione

Il problema di integrare le applicazioni esiste da sempre. Applicazioni scritte da soggetti diversi in tempi diversi non offrono strutture dati consistenti, non consentono l'esecuzione di processi trasversali, spesso non garantiscono una semantica comune e sono basate su tecnologie incompatibili. Le applicazioni sviluppate in passato erano spesso "mondi chiusi" e l'interazione con il mondo esterno era possibile solo attraverso un passaggio di dati fuori linea, estraendo i dati necessari per trasferirli in modo asincrono ad altre applicazioni all'interno della stessa azienda o di aziende diverse. Nel caso degli scambi di dati inter-aziendali, per ovviare alla complicazione del collegamento fra due sistemi remoti reciprocamente sconosciuti, sono stati sviluppati complessi standard di interscambio di documenti (EDI - *Electronic Document Interchange*), e

successivamente sistemi EDI "leggeri" basati sullo scambio di dati in formato XML. L'integrazione delle applicazioni si è evoluta negli ultimi anni per alcuni aspetti, tra cui la dimensione e la complessità delle applicazioni, lo scambio di dati a livello di ecosistemi e mercati e le tecnologie utilizzate. In passato, l'integrazione delle applicazioni operava su una scala più piccola, poiché si dovevano integrare applicazioni che oggi sono spesso già comprese in pacchetti omogenei come i sistemi ERP che inizialmente sembrava potessero eliminare all'origine il problema dell'integrazione. Invece il problema permane, ma su una scala maggiore, poiché è necessario integrare i sistemi ERP con applicazioni locali e altri sistemi come CRM (Customer Relationship Management), BI (Business Intelligence), SCM (Supply Chain Management), sistemi documentali e sistemi di automazione industriale. Contemporaneamente, la diffusione delle applicazioni basate su Internet ha accelerato l'esigenza di integrare le applicazioni tradizionali, spesso sviluppate in ambiente mainframe, con i servizi su Web, i sistemi di commercio elettronico e i portali aziendali.

Un altro importante tema della data integration è quello del *Master Data Management* (MDM). Ogni azienda ha un insieme di dati di importanza critica (Master Data), come prodotti, clienti e fornitori, che sono gestiti da varie applicazioni. Poiché la dispersione e la ridondanza dei master data provoca problemi di consistenza e di inefficienza nei processi aziendali, per risolvere questi problemi si possono spostare i Master Data al di fuori delle singole applicazioni. MDM è l'insieme di discipline, tecnologie e soluzioni in grado di creare e mantenere consistenti, aggiornati e completi i dati di importanza critica per l'azienda, fornendo una visione unica

ICT TREND: APPLICATION INTEGRATION

a utenti, applicazioni e processi. La realizzazione di una soluzione MDM richiede la raccolta di dati in formati diversi da molte fonti eterogenee, sia in modo schedulato sia in tempo reale, per alimentare il repository che contiene i Master Data.

Evoluzione degli strumenti

Inizialmente l'integrazione veniva affrontata costruendo ogni volta procedure di estrazione, filtraggio e importazione dei dati ad-hoc, mentre oggi è possibile utilizzare le potenti applicazioni general-purpose disponibili sul mercato. I primi prodotti nati per integrare applicazioni intra-aziendali (A2A), apparsi nel 1995, erano strumenti ETL (*Extract-Transform-Load*) usati per estrarre informazioni da una o più sorgenti, normalizzarle e inserirle in un'altra struttura dati. Alla fine degli anni '90 comparvero i primi prodotti in grado di integrare applicazioni inter-aziendali B2B attraverso Internet. Oggi esistono anche gli strumenti che permettono un'interazione diretta fra applicazioni e la vista in tempo reale federata (unificata) di dati eterogenei (EII – *Enterprise Information Integration*), appoggiandosi sui messaggi XML, sul middleware per lo scambio asincrono di messaggi (MOM) e su vari protocolli per l'interazione tra processi locali (come Microsoft COM, OMG CORBA), la condivisione distribuita tra ambienti omogenei (come IIOP, Java RMI e Microsoft DCOM), e la condivisione distribuita tra ambienti eterogenei attraverso standard Web aperti (come SOAP/XML o i web service, nell'ambito dell'architettura SOA). Un'altra forma di integrazione delle applicazioni concettualmente diversa dall'EAI

è il *Mashup*: la creazione di applicazioni composite, tipicamente applicazioni client "light" assemblate con funzioni esistenti per mezzo delle tecnologie Web come HTTP, RSS (Really Simple Syndication) e XML.

Enterprise Application Integration

Ci sono varie definizioni di *Enterprise Application Integration*. Per esempio, Wikipedia definisce EAI come "l'uso di software e principi architetturali di sistemi informatici per integrare un insieme di applicazioni informatiche aziendali", mentre per Gartner Group è la "condivisione senza restrizioni di dati e processi di business tra tutte le applicazioni e le fonti di dati nell'impresa".

I processi di Enterprise Application Integration devono collegare le applicazioni di un'organizzazione per semplificare e automatizzare al massimo i processi di business

Gli strumenti EAI possono essere utilizzati per vari scopi: assicurare la coerenza delle informazioni nei diversi sistemi (EII), creare viste federate di dati che risiedono in ambienti diversi (*data federation*) senza movimenti fisici di dati, estendere ed integrare i processi di business attraverso varie applicazioni, recuperare informazioni da applicazioni proprietarie, creare un'interfaccia utente comune unificata e coerente per le varie

applicazioni, per esempio tramite un portale aziendale che interagisce con diversi pacchetti software. L'obiettivo è comunque sempre quello di collegare dati, applicazioni e processi per realizzare vantaggi competitivi operativi ed economici. L'integrazione dei dati e delle applicazioni aziendali è una componente fondamentale della strategia globale di gestione del patrimonio informativo delle imprese (EIM – *Enterprise Information Management*).

Aspetti tecnici

I processi di Enterprise Application Integration devono collegare le applicazioni di un'organizzazione per semplificare e automatizzare al massimo i processi di business, senza apportare grandi modifiche alle applicazioni o alle strutture dei dati esistenti. È però necessario affrontare vari problemi pratici, poiché i vari sistemi che devono essere collegati spesso risiedono in differenti sistemi operativi, usano database e linguaggi di programmazione differenti, e talvolta sono applicazioni legacy non documentate e non supportate. Poiché al crescere del numero delle applicazioni il numero delle connessioni point-to-point cresce in modo esponenziale, è necessario definire e implementare un bus software (*Enterprise Service Bus*) a cui si connettono tutte le applicazioni tramite opportuni moduli di interfaccia (adattatori), che devono essere in grado di supportare tutti gli ambienti da collegare: database relazionali, database legacy, vari formati di file, applicazioni proprietarie, XML, formati standard di messaggistica come EDI, SWIFT e Health Level Seven (HL7), code di messaggi, e-mail, siti Web, strumenti office e sistemi

ICT TREND: APPLICATION INTEGRATION

documentali. Inoltre, gli strumenti EAI devono essere in grado di supportare il maggior numero possibile di piattaforme hardware e software, come mainframe (IBM z/OS e z/Linux), sistemi midrange (IBM System i – AS/400), sistemi Unix, Linux e Windows.

I vantaggi offerti dagli strumenti EAI sono la possibilità di accedere alle informazioni di sistemi diversi in tempo reale o differito, di aumentare l'efficienza dei processi di business e dell'organizzazione, mantenendo l'integrità e l'univocità delle informazioni, semplificando al contempo lo sviluppo e la manutenzione. Gli svantaggi sono rappresentati dai notevoli

costi degli strumenti, dallo sforzo iniziale richiesto per formare i tecnici sugli strumenti, implementare i sistemi EAI in termini di tempi e di risorse, e infine dal pericolo di dipendere da un singolo fornitore se le tecnologie di collegamento utilizzate non sono sufficientemente aperte.

I prodotti disponibili

Gli strumenti per estrazione, trasformazione e caricamento (ETL) di prima generazione, usati prevalentemente per alimentare warehouse e data mart, si sono estesi fino a combinare la data integration, la gestione dei metadati e l'interpretazione se-

mantica. Non esiste ancora sul mercato una singola piattaforma o suite EAI in grado di soddisfare qualsiasi esigenza.

Gartner Group stima una dimensione del mercato mondiale degli strumenti per l'integrazione dei dati pari a \$1,44 miliardi di dollari alla fine del 2007, e prevede un tasso di crescita annuale superiore al 17% fino al 2012.

Nel seguito sono elencati (in ordine alfabetico) i principali fornitori riportati nel Magic Quadrant Gartner Group degli strumenti EAI (2008).

IBM è uno dei principali attori, con notevoli capacità di data integration, anche grazie all'ac-

docPorter. Per un ufficio senza carta.



Michele Balduani
ALA Informatica

accedi al Web Seminar su:
www.cbritaly.it/article.php?id=597

L'ufficio è più leggero con DocPorter, il sistema di Gestione Documentale, Business Process Management, fatturazione elettronica e conservazione sostitutiva per la gestione efficiente e trasparente dei documenti.

I rapporti con i clienti, con i fornitori ed i processi interni si basano sulla gestione dei documenti scambiati tra le parti; è importante disporre di un sistema centralizzato di raccolta e di ricerca rapida delle informazioni che risponda agli standard di sicurezza.

Migliorare i processi aziendali, sveltendo le procedure ed eliminando le discontinuità operative determinate dall'utilizzo dei supporti cartacei, sarà la sfida del prossimo futuro.

Con docPorter BPM generate rapidamente i modelli dei flussi di lavoro della vostra organizzazione, integrando documenti, sistemi e persone.

ALA
informatica
www.alainformatica.com

ICT TREND: APPLICATION INTEGRATION

quisizione di Cognos. Il sistema Information Server costituisce una delle più complete piattaforme di data integration.

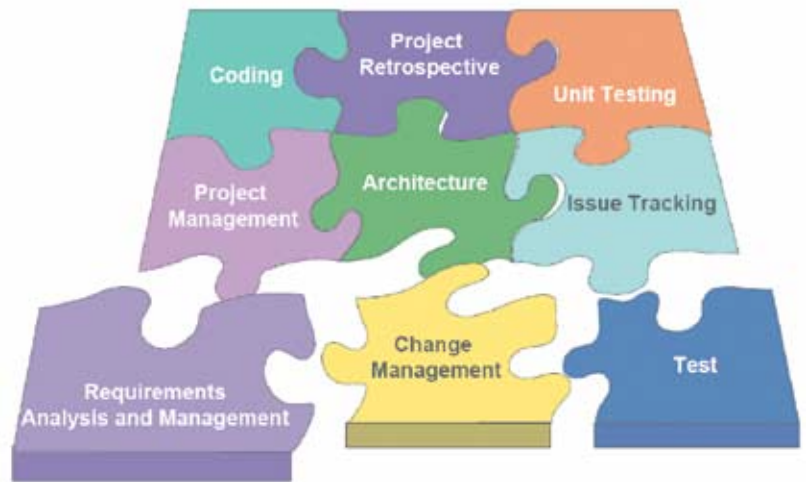
Informatica (The Data Integration Company) è il maggiore fornitore specializzato di questo mercato, con una base installata di oltre 3.000 grandi clienti, numerosi progetti presso le più importanti aziende nel mondo e il maggior numero di specialisti EAI. PowerCenter è una piattaforma unificata per l'integrazione dei dati a livello enterprise, in grado di accedere ai dati di qualunque sistema e in ogni formato, che grazie a una notevole scalabilità può essere utilizzato come base per tutti i progetti di integrazione dati e per le attività di integrazione aziendali.

Microsoft offre SQL Server Integration Services (SSIS), che supporta massicci trasferimenti di dati via ETL e si interfaccia con BizTalk Server. SQL Server 2008 espande la connettività SSIS a fonti di dati eterogenee, come SAP, Oracle, Teradata e IBM DB2.

Open Text realizza l'integrazione di dati e contenuti con il proprio prodotto Genio (sviluppato da Hummingbird, acquisita nel 2006), uno strumento ETL di semplice uso, spesso usato a livello dipartimentale.

Oracle offre due strumenti specifici: Oracle Warehouse Builder (OWB), compreso nelle licenze dei DBMS Oracle, e Oracle Data Integrator (ODI), che appartiene alla famiglia di prodotti Middleware Oracle Fusion. Con l'acquisizione di Business Objects, SAP ha arricchito la propria offerta di data integration, che oltre agli strumenti SAP include ora BusinessObjects Data Federator e la piattaforma Data Services.

SAS offre Enterprise Data Inte-



Il mercato degli strumenti di Enterprise Applications Integration costituisce quindi una nicchia in rapida crescita, intimamente collegata all'evoluzione delle Service Oriented Architecture

gration Server, con funzioni ETL e di data federation, che funziona con tutti i maggiori sistemi operativi ed è in grado di collegarsi a tutti i tipi di dati.

Gli strumenti EAI di Sun Microsystems sono parte della piattaforma Java Composite Application Platform Suite (CAPS), che comprende tutti i componenti necessari per implementare un'architettura SOA.

Sybase ha diversi prodotti per realizzare vari stili di data delivery: ETL, data federation e replica. Gli strumenti principali sono Sybase Replication Server, che supporta la replica di database eterogenei, e PowerDesigner per la modellazione di dati e metadati.

Tibco Software offre diverse soluzioni EAI basate su un'infrastruttura SOA comune: integrazione delle applicazioni message-oriented con BusinessWorks,

ETL con Data Exchange e Collaborative Information Manager per le applicazioni MDM.

Esistono anche progetti e prodotti Open Source, soprattutto sulla piattaforma Apache (ActiveMQ, Camel, ServiceMix, Synapse), anche nella versione integrata, certificata e supportata dalla distribuzione della comunità FUSE (Message Broker, Mediation Router, ESB).

MuleSource offre un software di infrastruttura SOA Open Source costituito dall'enterprise service bus (ESB) e piattaforma di integrazione Mule.

Il mercato degli strumenti EAI costituisce quindi una nicchia in rapida crescita, intimamente collegata all'evoluzione delle architetture SOA, che partendo dalle grandi organizzazioni si diffonderà sempre di più anche tra le medie imprese. **B**