



IL DATALINEAGE LA DATAVIRTUALIZATION PER SEMPLIFICARE LA DISTRIBUZIONE DEL DATO, DA UN DB CENTRALE AI DIVERSI REPARTI NEL SETTORE BANCARIO

IL CLIENTE

Il cliente per cui abbiamo sviluppato il progetto è nel settore bancario.

Nello specifico, l'area per cui abbiamo sviluppato il progetto ha lo scopo di fornire dati alle varie aree dell'azienda (ogni area ha necessità di visualizzare dati differenti: ad esempio i responsabili marketing avranno bisogno di un tipo di dati differenti dai dati utilizzati dal personale allo sportello)

PROBLEMATICHE ED ESIGENZE

Il reparto IT del cliente presenta alcuni problemi a inviare i dati nella visualizzazione più appropriata alle varie applicazioni/utenti. Una delle esigenze primarie è garantire la certificazione del dato.

Il database da cui estraggono i dati (una specie di Data Warehouse, utilizzato per fornire i dati più alle applicazioni in modo automatico che agli utenti) è molto strutturato e modificarlo renderebbe molto complessa la certificazione del dato che per una banca è fondamentale.

Prima del progetto per ottenere i dati adatti ad ogni applicazione venivano create delle query complesse che andavano direttamente ad interrogare il database.

Alcuni svantaggi:

- Complessità nella scrittura delle query/flussi
- Complessità gestione dei flussi/query
- In caso cambino delle esigenze di lettura/scrittura per un'applicazione è molto complesso ed è necessario ritoccare tutto il flusso (anche per un minimo cambiamento) - questo comporta rallentamento nelle risposte da dare agli applicativi - aumento continuo della complessità di erogazione dei dati e maggiore forza lavoro impegnata su questi flussi
- Per mantenere il dato certificato, il database di partenza non può essere semplificato, potrebbe essere replicato con lo svantaggio di occupare il doppio dello spazio e avere due ambienti diversi che dovrebbero essere aggiornati simultaneamente per essere allineati.

Risolvere il problema attraverso la replicazione del dato comporterebbe un aumento dei costi di infrastruttura e del livello di complessità (basti pensare che una modifica su un campo di origine costringerebbe i tecnici ad identificare il tracciato dal campo origine - allo staging - ad DB complesso - alla query complessa, e tutto ciò per ogni singolo reparto a cui viene fornito quel dato)

La modifica di un campo sorgente costringe i tecnici ad individuare tutti gli applicativi ai quali è stato fornito il dato intervenendo e riportando le modifiche su tutti i flussi che li alimentano.

SOLUZIONE

Per far fronte a questo problema è stato utilizzato Denodo, strumento di Data Virtualization, che permette di virtualizzare il dato, così da semplificarlo e fornirlo alle applicazioni attraverso query semplici.

Vantaggi data lineage: mappatura del campo - permette di capire per ogni campo quali e quanti applicativi alimenta - quindi nel caso in cui venga fatta una modifica, riesco a rintracciare su quali applicativi/ si ripercuote l'eventuale modifica del campo.

Denodo ha un suo data lineage per delineare tutto il percorso del dato e quale impatto ha la sua modifica sul processo globale. Grazie al data lineage è possibile individuare in tempi brevi, in modo grafico e semplice dove è necessario modificare il campo.

Inoltre, il data lineage di Denodo può essere integrato con diversi data lineage sul mercato, così da lasciare un solo strumento agli utenti per avere tutta la traccia del dato. Denodo sposta la cache su sistemi esterni, così facendo il cliente può mantenere solo la cache su sistemi che ha già "in casa" senza appesantire il motore di Data Virtualization.

Denodo inoltre è integrato con i prodotti MDM (master data management¹) leader sul mercato così da importare la struttura delle tabelle e le relazioni senza doverla rifare da zero.

¹ Master data management: Master data management (MDM) è una disciplina abilitata alla tecnologia in cui business e IT collaborano per garantire uniformità, accuratezza, gestione, coerenza semantica e responsabilità delle risorse di dati master condivise ufficiali dell'azienda. I dati anagrafici sono l'insieme coerente e uniforme di identificatori e attributi estesi che descrive le entità principali dell'azienda, inclusi clienti, prospecti, cittadini, fornitori, siti, gerarchie e piano dei conti. (fonte Gartner)