

Sistemi Informativi Distribuiti

In dipendenza del luogo dove è posta fisicamente la base di dati, i sistemi informativi si dividono in sistemi **centralizzati** (quando la base di dati è memorizzata in un unico elaboratore) e sistemi **distribuiti** (quando la base di dati è memorizzata in due oppure più elaboratori collegati in rete locale o geografica).

In base alla modalità di funzionamento, i sistemi centralizzati si distinguono in sistemi **monoutente**, in cui i servizi sono finalizzati a una sola persona per volta, o **multiutente**, quando sia la base di dati sia i programmi applicativi di gestione sono posti su un server di database accessibile dai computer client mediante una rete di computer.

La costruzione di un sistema informativo distribuito avviene generalizzando quella relativa al caso centralizzato, con la differenza che la base dei dati è sostituita con un database distribuito (la memorizzazione dei dati posti avviene in luoghi geografici diversi) e le operazioni sono realizzate mediante interrogazioni decentrate (le interrogazioni coinvolgono più server della rete).

Gli ambienti software per la realizzazione di database e servizi decentrati sono indicati con il termine **DDBMS** (Distributed DataBase Management System).

Organizzazione di un Sistema Informativo Distribuito

Tra le diverse architetture decentrate, possiamo riconoscere i sistemi informativi distribuiti **propriamente detti**, **multidatabase** e con una **organizzazione centrale-periferica**:

1. I sistemi informativi distribuiti "propriamente detti" impiegano un modello dei dati, nel quale le tabelle sono frammentate e/o duplicate nei server di una rete di computer e tecniche di interrogazione distribuita.
2. I sistemi multidatabase (distributed multidatabase system) sono costituiti da un insieme di sistemi informativi locali perfettamente autonomi collegati tra loro mediante una rete di computer.
3. Un sistema distribuito centrale-periferico è formato da un insieme di sottosistemi informativi periferici (branch system) e un sottosistema centrale (corporate system) in comunicazione tra loro mediante una rete di computer locale o geografica. La sede centrale

dispone del server primario e del database centrale e i computer delle sedi periferiche hanno il compito di archiviare temporaneamente nuove registrazioni su basi di dati locali, per poi aggiornare periodicamente il database centrale con le nuove transazioni.

Basi di dati distribuite

Con il termine base di dati distribuita (DDB: Distributed DataBase) si intende un insieme di dati memorizzati in modo permanente su supporti fisici di almeno due server differenti (in comunicazione mediante una rete di computer), strutturati sulla base di uno schema o modello, protetti dall'esterno (sicurezza) e affidabili, disponibili per essere utilizzati da applicazioni e utenti diversi nella rete di calcolatori, indipendentemente dal luogo geografico dove i dati sono memorizzati (proprietà di trasparenza).

Per realizzare il modello relazionale di un database distribuito occorre progettare lo schema relazionale completo, come se la base di dati fosse centralizzata, frammentare le tabelle del database e/o duplicare i dati, assegnare le tabelle frammentate e/o duplicate ai diversi server di database della rete di computer e infine creare (se necessari) degli schemi esterni (viste) per i diversi utenti del sistema.

I dati in un database distribuito sono memorizzati in server di una rete di computer. Le tabelle del modello relazionale sono quindi frammentate (divise in parti) seguendo tre tecniche diverse denominate frammentazione **verticale**, **orizzontale** o **mista** (ibrida).

1. Nella frammentazione verticale (o per colonne), le tabelle sono divise in colonne (attributi) e memorizzate in server differenti nella rete.
2. Nella frammentazione orizzontale (o per righe), le tabelle sono divise per righe, per poi allocare i frammenti ottenuti sulle memorie di massa di differenti server della rete.
3. Nella frammentazione mista (o ibrida), le tabelle sono suddivise nei diversi server della rete applicando, in sequenza, prima la frammentazione orizzontale e in seguito quella verticale.

La duplicazione dei dati è necessaria nei casi in cui si devono realizzare interrogazioni (query) distribuite in tempo reale, evitando i tempi di trasmissione sulla rete.

DDBMS

Un DDBMS (**D**istributed **DB**MS) è un ambiente software applicativo complesso per la gestione di più basi di dati decentralizzate, il quale mette a disposizione del gruppo dei programmatori un linguaggio interno, in cui sono presenti i sottoinsiemi di istruzioni DDL, DML e DCL o uno oppure più linguaggi di programmazione esterni.

Gli obiettivi sono quelli di creare un ambiente integrato di lavoro per la realizzazione delle parti che compongono un sistema informativo distribuito automatizzato, di attuare tutte le proprietà di una base di dati distribuita quali sicurezza, affidabilità, indipendenza fisica e logica dei dati e di realizzare la proprietà di trasparenza.

Gestione delle transazioni

In un sistema informativo centralizzato multiutente oppure distribuito, uno dei problemi principali da gestire per i DDBMS è l'accesso simultaneo (concorrente) di due oppure più utenti, che vogliono eseguire operazioni sui medesimi dati, su un database memorizzato in un server. I DDBMS fondano la gestione delle operazioni concorrenti sui database sul concetto di transazione.

Una transazione (transaction) è un insieme di una oppure più istruzioni del linguaggio interno trattate dal DDBMS come una singola unità di elaborazione.

Una transazione può essere implicita (se creata automaticamente dal DDBMS nei casi di modifiche dello stato del database) oppure esplicita (se definita dal programmatore).

Ogni transazione, prima di essere eseguita, è registrata nel registro (lock) delle transazioni, memorizzato in un file distinto da quello principale del database.

Quando una transazione sta utilizzando dei dati condivisi, li pone in uno stato di blocco (lock), per evitare che altre operazioni intervengano in simultanea sulla medesima risorsa.