

CAP4

SISTEMI

CENTRALIZZATI E

DISTRIBUITI

Ing. Daniele Corti



copyright

all rights reserved

Copyright © Ing. Daniele Corti 2013

www.ingdanielecorti.it

Tutti i diritti sono riservati a norma di legge e a norma delle convenzioni internazionali.

Ver.1.0

PREREQUISITI

- ✓ Conoscenze sulle tipologie di archivio.
- ✓ Conoscenze sul concetto di base di dati (database).

OBIETTIVI

- ✓ Riconoscere le differenze fra archivi centralizzati e distribuiti.

ARGOMENTI

- ✓ Sistemi centralizzati e distribuiti in relazione al luogo dove è posta fisicamente la base di dati.
- ✓ Modello dei dati.
- ✓ Interrogazione distribuita.
- ✓ Esempio

CAP4 - Sistemi centralizzati e distribuiti

In relazione al luogo dove è posta fisicamente la base di dati

I sistemi informativi, sulla base del luogo dove il DB è realmente dislocato, si possono suddividere in:

- **SISTEMI CENTRALIZZATI**
- **SISTEMI DISTRIBUITI**

SISTEMA CENTRALIZZATO

Un sistema si dice centralizzato quando il DB è memorizzato in un unico elaboratore.

Il DB può essere **locale**, residente sullo stesso calcolatore dell'utente, o **remoto** se è residente su un altro computer interno o esterno alla azienda (**server**).

In base alla modalità di funzionamento, i sistemi centralizzati si distinguono in **Sistemi monoutenti**, 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 DB accessibile dai computer client mediante una rete di computer (**architettura client-server**).

SISTEMA DISTRIBUITO

Un sistema si dice distribuito quando il DB è memorizzato in due o più elaboratori collegati in rete locale o geografica.

Il database può essere dislocato in più computer situati nello stesso luogo, oppure distribuito in una rete di computer connessi tra loro sotto forma appunto di sistema distribuito.

Si tratta di basi di dati con informazioni costituenti una sola struttura logica ma fisicamente memorizzate su diversi elaboratori autonomi e connessi attraverso una rete. La localizzazione dei dati è trasparente alle applicazioni degli utenti. Le informazioni sono disponibili per essere utilizzate da applicazioni e utenti diversi nella rete di calcolatori, indipendentemente dal luogo geografico dove i dati sono memorizzati.

I database distribuiti sono nati per risolvere esigenze di enti, organizzazioni o aziende fortemente decentrate con la necessità di manipolare grosse quantità di dati.

Gli aspetti umani, organizzativi ed economici dell'ente e degli utilizzatori dei dati traggono molto vantaggio da un'organizzazione di questo tipo; purtroppo, esistono anche alcuni problemi (di natura tecnica):

- Decentralizzazione del controllo
- Problemi di comunicazione
- Complessità del software di gestione della base di dati (DDBMS).

I database distribuiti si basano sul concetto di Client/Server.

~~In funzione del tempo entro cui deve avvenire la risposta rispetto alla richiesta di un utente i sistemi informativi distribuiti possono classificarsi in:~~

- ~~• OLTP (Elaborazioni di Transazioni On Line) nei quali le operazioni devono essere svolte nel minor tempo possibile. Esempi: attraverso Internet, le operazioni di acquisto, le transazioni bancarie, le operazioni in borsa, le prenotazioni per treni e aerei.~~
- ~~• OLBP (Elaborazioni Batch On Line) nei quali i dati delle operazioni, raccolti e memorizzati precedentemente, sono periodicamente elaborati per aggiornare i dati nei file del database. Esempio: le buste paghe mensili dei dipendenti di un'azienda a livello nazionale. Tutto il database anziché essere memorizzato su un unico elaboratore, viene memorizzato su diversi elaboratori, detti server di database (della rete). I sistemi distribuiti utilizzano~~
- ~~• Un opportuno modello di dati.~~
- ~~• Tecniche di interrogazione distribuita.~~

MODELLO DEI DATI

Rispetto al caso centralizzato, gli schemi del database sono resi più complicati dalla necessità di distribuire fisicamente i dati su più computer.

Per questo motivo le tabelle sono frammentate: le tabelle cioè vengono suddivise in più righe e/o colonne ed ciascuna parte è memorizzata su server differenti.

Nei sistemi distribuiti, tra le cause che possono contribuire a rendere poco affidabili le informazioni vi sono i guasti HW/SW nei server di database e gli errori nella rete di telecomunicazioni: se per una di queste cause un server non è disponibile, allora viene a mancare una parte di database. Si devono, quindi, realizzare meccanismi per evitare la perdita dei dati; un modo è quello di creare tabelle duplicate nel server di dati. I dati pertanto vengono completamente o parzialmente duplicati in più server che automaticamente ed autonomamente si aggiornano l'un l'altro.

Come si costruisce il modello di database distribuito.

1. Progetto dello schema relazionale completo, come se la base fosse centralizzata;
2. Frammentazione delle tabelle del database e/o duplicazione dei dati;
3. La frammentazione può essere verticale o orizzontale. Nel primo caso le tabelle sono divise in colonne (la chiave primaria è sempre presente) e memorizzate in server differenti; nel secondo caso, le tabelle sono divise per righe.
4. Assegnazione delle tabelle frammentate e/o duplicate ai diversi server;
5. Creazione delle viste logiche per i diversi utenti del sistema.

INTERROGAZIONE DISTRIBUITA

Un'interrogazione distribuita, in un DDBMS, viene eseguita con la seguente procedura:

Da un computer viene effettuata la richiesta di alcune informazioni ad un server di database.

1. localizzazione dei dati: il server (sulla base delle proprie conoscenze sul database relazionale complessivo) deve individuare in quali server sono allocati i frammenti delle tabelle, contenenti le informazioni richieste;
2. decomposizione dell'interrogazione: il server divide la query globale in altre parziali (una per ogni server interessato);
3. richiesta dei dati: ad ogni server interessato viene chiesto di eseguire la query corrispondente;
4. esecuzione delle query locali: ogni server (anche contemporaneamente agli altri) esegue la propria query;
5. raccolta dei dati parziali: ogni server invia il risultato della propria query al server a cui è stata richiesta l'esecuzione della query globale;
6. elaborazione e presentazione dei risultati finali: il server che ha ricevuto la richiesta globale raccoglie tutti i dati, li elabora e fornisce all'applicazione client le informazioni richieste.

ESEMPIO DI DATABASE DISTRIBUITO

Consideriamo il seguente progetto di sistema informativo distribuito.

Un'azienda vuole creare un fondo pensione comune organizzando sul territorio nazionale delle sedi dove i risparmiatori possono versare annualmente le somme dovute. Questi versamenti, nel corso degli anni, possono essere effettuati presso sedi diverse, perché, ad esempio, i risparmiatori hanno cambiato il luogo di residenza. Si supponga, per semplicità, di volere gestire solo tre sedi così decentrate:

una sede a, Milano per il nord Italia;

una sede a Roma per il centro;

una sede a Napoli per il sud e le isole.

Al termine del piano previdenziale ogni risparmiatore chiede la ricongiunzione dei versamenti presso una delle filiali, che funzionerà da sede accentratrice per la liquidazione in un'unica soluzione del montante dei suoi versamenti.

La sede accentratrice potrebbe anche essere diversa dalle quelle che hanno ricevuto i versamenti. Il servizio di "ricongiunzione dei versamenti" deve anche essere disponibile via Intemel, in modo che ogni cliente autorizzato sia in grado di accedere alla sede centrale dalla propria abitazione con un browser via MODEM.

Applichiamo il metodo appena descritto per progettare un modello relazionale distribuito del sistema informativo di supporto del fondo pensione.

Lo schema relazionale, senza tenere conto della distribuzione dei dati, è formato dalle seguenti tabelle:

Risparmiatori (Codice Risparmiatore, Cognome, Nome, Indirizzo, Data di Nascita, Codice Fiscale)

Versamenti (Codice Risparmiatore, Anno Versamento, Codice Filiale, Somma Versata)

Filiali (Codice Filiale, Sede)

La situazione, dal punto di vista della rete di calcolatori e della distribuzione dei dati del database, è riportata schematicamente nella seguente figura.

In essa si nota che:

§ in ogni sede è presente una rete LAN, composta dai terminali Client, un Server di database per contenere i dati, da un Server di rete per la connessione con le altre sedi e da un Modem per realizzare la connessione remota;

§ le tre sedi sono collegate tra loro, per mezzo dei Modem, realizzando così una Intranet geografica;

Borse di studio: come ottenerle

Trucchi per studiare meglio

Libri usati: mappa dei mercatini

- nella LAN della sede di Roma è presente un Server per la produzione del sito WEB, che permetterà agli iscritti al fondo pensione dotati di autorizzazione di poter verificare la propria situazione direttamente dal computer della propria abitazione (fornendo questa possibilità, la rete aziendale diventa una Extranet);

- per garantire l'accesso a Internet, l'azienda si appoggia a un server Web di un ISP (Internet Service Provider). Le pagine Web sono realizzate localmente presso la sede di Roma e trasferite periodicamente sul server dell'ISP.

Il contenuto delle tabelle per ciascuna sede è il seguente

SEDE DI ROMA

TABELLA VERSAMENTI (frammento)

Codice Risparmiatore	Anno Versamento	Codice Filiale	Somma Versata
R-001	2000	F-001	1250,00
R-001	2001	F-001	1250,00
R-002	2002	F-001	1250,00
R-001	2002	F-001	1250,00
R-002	2003	F-001	1250,00
R-001	2003	F-001	1250,00
R-002	2004	F-001	1250,00
R-003	2009	F-001	1250,00

TABELLA FILIALI

Codice Filiale	Sede
F-001	ROMA
F-002	MILANO
F-003	NAPOLI

TABELLA RISPARMIATORI

Codice Risparmiatore	Cognome	Nome	Indirizzo	Data di Nascita
R-001	ALLIEVI	FRANCESCA	VIA DEL CORSO, 44 - MILANO	10/10/1950
R-002	MISEROTTI	MARCO	CORSO SALERNO, 56 – NAPOLI	11/01/1951
R-003	SCARDELLINI	LUISA	CORSO STATI UNITI, 55/A - ROMA	10/01/1955

SEDE DI MILANO

TABELLA VERSAMENTI (frammento)

Codice Risparmiatore	Anno Versamento	Codice Filiale	Somma Versata
----------------------	-----------------	----------------	---------------

R-003 2005 F-002 1250,00
R-002 2005 F-002 1250,00
R-003 2006 F-002 1250,00
R-002 2006 F-002 1250,00
R-003 2007 F-002 1250,00
R-003 2008 F-002 1250,00
R-001 2008 F-002 1250,00
R-001 2009 F-002 1250,00

SEDE DI NAPOLI

TABELLA VERSAMENTI (frammento)

Codice Risparmiatore	Anno	Versamento	Codice Filiale	Somma Versata
R-003	2002	F-003	1250,00	
R-003	2003	F-003	1250,00	
R-003	2004	F-003	1250,00	
R-001	2004	F-003	1250,00	
R-001	2005	F-003	1250,00	
R-001	2006	F-003	1250,00	
R-002	2007	F-003	1250,00	
R-001	2007	F-003	1250,00	
R-002	2008	F-003	1250,00	
R-002	2009	F-003	1250,00	

Il risparmiatore chiede, presso la sede di Milano, la ricongiunzione delle somme versate.

Il computer dell'operatore effettuerà la richiesta al proprio server (di Milano). Il server di Milano dovrà richiedere la trasmissione:

§ delle somme versate nei vari anni ai server remoti delle sedi di Roma e Napoli e a quello locale di Milano;

§ dei dati personali al server della sede di Roma.

Query Richiesta dei dati anagrafici (al server di Roma, da parte del server di Milano)


```
SELECT [Codice Risparmiatore], Cognome, Nome, Indirizzo, [Data di Nascita]
FROM ServerROMA.Risparmiatori
WHERE [Codice Risparmiatore] = "R-001"
```

1^ Query Parziale Versamenti (richiesta dal Server di Milano al server di Roma, tramite una connessione remota)

```
SELECT [Codice Risparmiatore], [Anno Versamento], [Codice Filiale], [Somma Versata]
FROM ServerROMA.Versamenti
WHERE ([Codice Risparmiatore] = "R-001" AND ([Codice Filiale] = "F-001"))
```

2^ Query Parziale Versamenti (richiesta dal Server di Milano al server di Napoli, tramite una connessione remota)

```
SELECT [Codice Risparmiatore], [Anno Versamento], [Codice Filiale], [Somma Versata]
FROM ServerNAPOLI.Versamenti
WHERE ([Codice Risparmiatore] = "R-001" AND ([Codice Filiale] = "F-003"))
```

3^ Query Parziale Versamenti (eseguita localmente sul Server di Milano)

```
SELECT [Codice Risparmiatore], [Anno Versamento], [Codice Filiale], [Somma Versata]
FROM ServerMILANO.Versamenti
WHERE ([Codice Risparmiatore] = "R-001" AND ([Codice Filiale] = "F-002"))
```

La Query complessiva dei versamenti è formata dall'Unione delle tre query parziali

RISULTATI OTTENUTI DALLE QUERY

Dal Server di Roma (query remota)

Codice Risparmiatore	Cognome	Nome	Indirizzo	Data di Nascita
R-001	ALLIEVI	FRANCESCA	VIA DEL CORSO, 44 - MILANO	10/10/1950

Codice Risparmiatore	Anno Versamento	Codice Filiale	Somma Versata
R-001	2000	F-001	1250,00
R-001	2001	F-001	1250,00
R-001	2002	F-001	1250,00
R-001	2003	F-001	1250,00

Dal Server di Milano (query locale)

Codice Risparmiatore Anno Versamento Codice Filiale Somma Versata

R-001 2008 F-002 1250,00

R-001 2009 F-002 1250,00

Dal Server di Napoli (query remota)

Codice Risparmiatore Anno Versamento Codice Filiale Somma Versata

R-001 2004 F-003 1250,00

R-001 2005 F-003 1250,00

R-001 2006 F-003 1250,00

R-001 2007 F-003 1250,00

Il Server di Milano unisce i risultati

Dati personali

Codice Risparmiatore Cognome Nome Indirizzo Data di Nascita

R-001 ALLIEVI FRANCESCA VIA DEL CORSO, 44 – MILA

NO 10/10/1950