

Basi di Dati Parallele

Capitolo 3

Basi di dati – Architetture e linee di evoluzione
P. Atzeni, S. Ceri, P. Fraternali, S. Paraboschi, R. Torlone



1

Scalabilità delle applicazioni

- Carico
 - insieme di tutte le applicazioni (query)
- Scalabilità
 - abilità di conservare prestazioni elevate al crescere del carico
- Dimensioni di crescita
 - numero delle query
 - complessità delle query

2

● ● ● Tipologie di carico

- Transazionale
 - carico: transazioni brevi
 - misura: tps (transazioni al secondo)
 - tempo di risposta: pochi secondi
- Analisi dei dati
 - carico: query SQL complesse
 - tempo di risposta: variabile

3

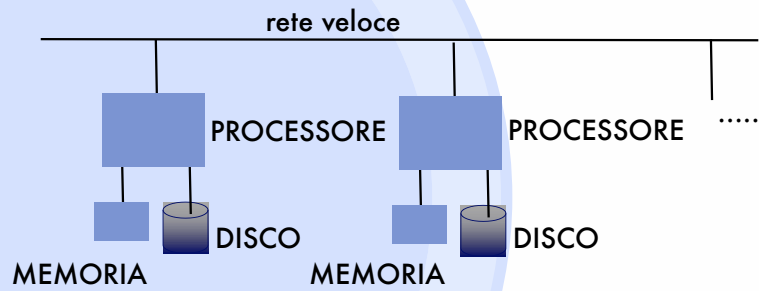
● ● ● Parallelismo

- Ottenuto tramite molti processori che cooperano in un'unica architettura informatica
- Due tipi di **parallelismo**
 - **Inter-query**
ciascuna query affidata ad un solo processore
(per carichi transazionali)
 - **Intra-query**
ciascuna query affidata a molti processori
(per carichi di analisi dati)

4

Architetture a confronto

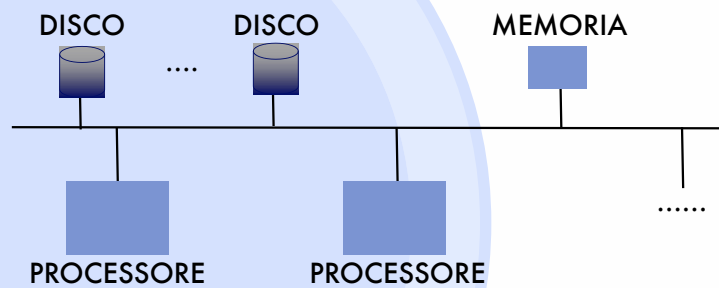
SHARED-NOTHING



5

Architetture a confronto

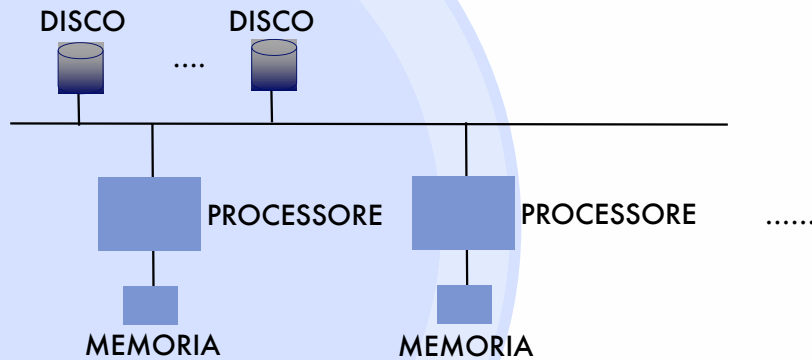
SHARED-MEMORY



6

Architetture a confronto

● SHARED-DISK



7

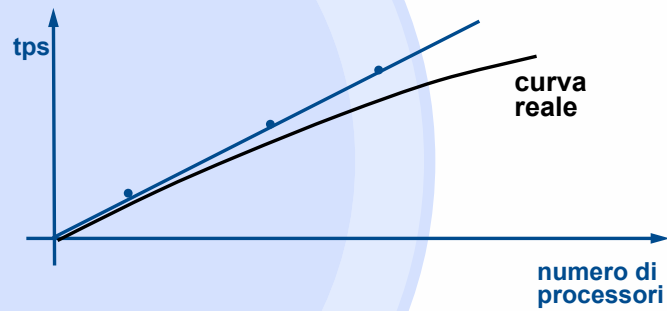
Benchmark

- Metodi per confrontare le prestazioni di sistemi diversi (in competizione)
- Standardizzazione
 - del database
 - del carico
 - del codice delle transazioni
 - modalità di invio
 - frequenza di arrivo
 - della modalità di misurazione
- Varie tipologie di carico (TPC-C, TPC-H, TPC-W)

8

●●● Curva di speed-up

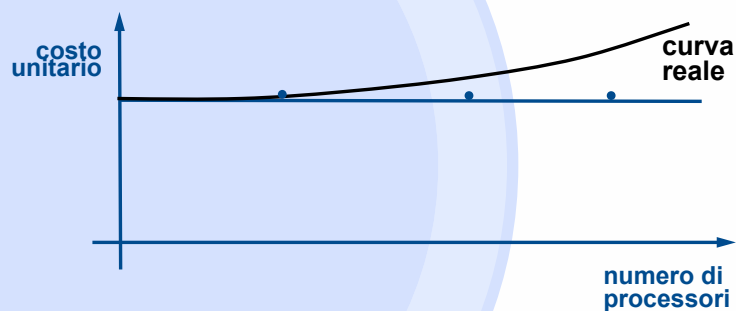
- Misura il crescere di efficienza al crescere del numero di processori



9

●●● Curva di scale-up

- Misura il crescere di costo unitario complessivo per transazione al crescere del numero di processori



10

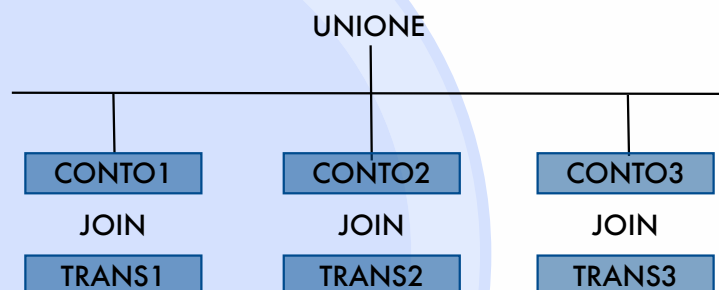
○ ● ● Join distribuito

- E' l'operazione di analisi dei dati più onerosa
- Consideriamo:



11

○ ● ● Join distribuito



12

Requisiti per il join distribuito

- I domini degli attributi di join devono essere partizionati e ogni partizione assegnata ad una coppia di frammenti
 - ad esempio su valori numerici tra 1 e 300000:
 - da 1 a 100000
 - da 100001 a 200000
 - da 200001 a 300000
- In molti sistemi paralleli i dati vengono inizialmente ridistribuiti sui dischi per ottenere questa distribuzione

13

Basi di Dati Replicate

Capitolo 3

Basi di dati - Architetture e linee di evoluzione
P. Atzeni, S. Ceri, P. Fraternali, S. Paraboschi, R. Torlone



14

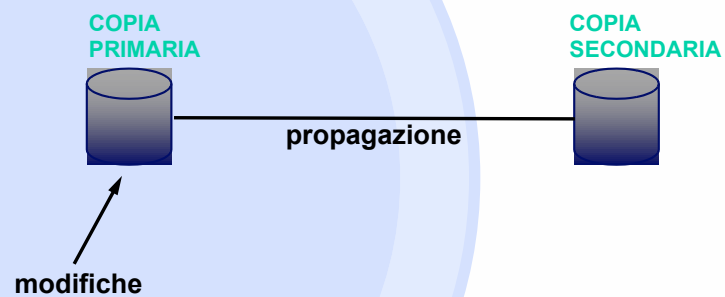
● ● ● Replicazione dei dati

- E' un ingrediente fondamentale dei sistemi informativi
- Motivazioni:
 - efficienza
 - affidabilità
 - autonomia

15

● ● ● Modalità di replicazione

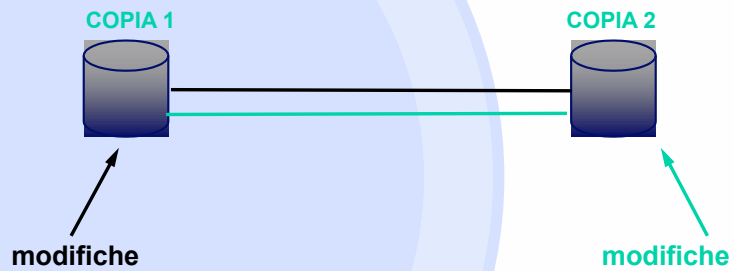
● **Asimmetrica**



16

○ ● ● Modalità di replicazione

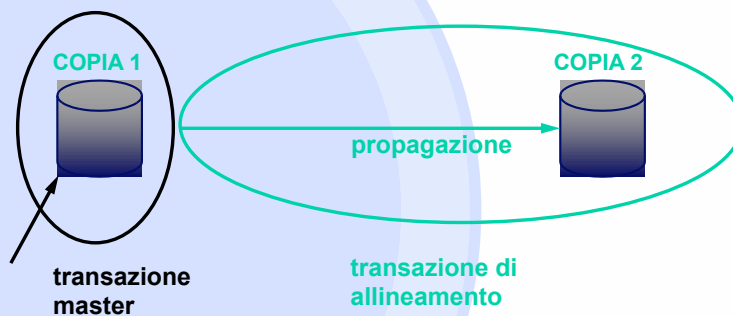
● Simmetrica



17

○ ● ● Modalità trasmissione variazioni

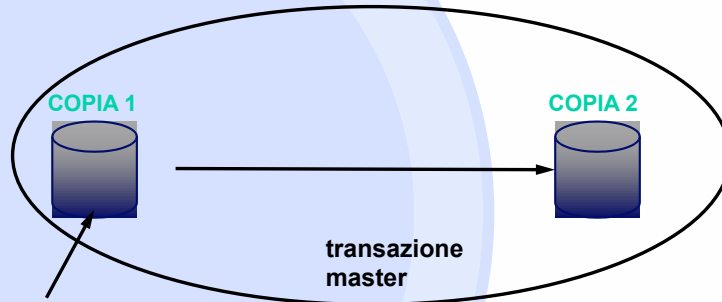
● Trasmissione asincrona



18

○ ● ● Modalità trasmissione variazioni

● **Trasmissione sincrona**



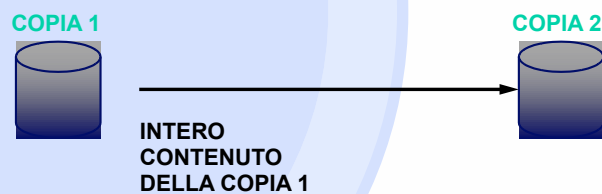
19

○ ● ● Modalità allineamento

● **Refresh**

● **Allineamento:**

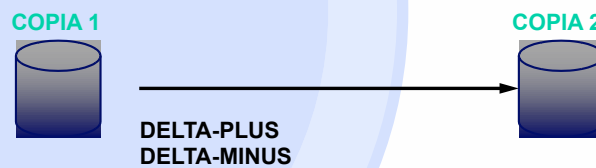
- periodico
- a comando
- ad accumulo di variazione



20

● ● ● Modalità allineamento

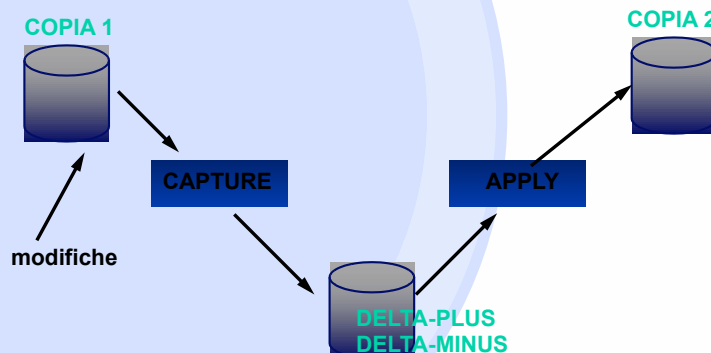
- **Incrementale**
- **Allineamento:**
 - periodico
 - a comando
 - ad accumulo di variazione



21

● ● ● Meccanismi per la replicazione

- **Asimmetrica, asincrona, incrementale**
- **Prodotto: *Replication Manager***
 - modulo CAPTURE
 - modulo APPLY



22

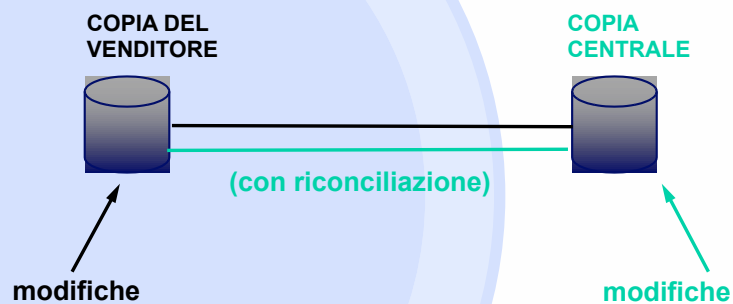
○ ● ● Replicazione in computer mobili

- Computer mobili
 - saltuariamente collegati ad una rete
- Copie disconnesse per ore o giorni interi, poi ricollegate (riconciliazione)
- Applicazione
 - agenti di vendita mobili

23

○ ● ● Allineamento di copie disconnesse

- Richiede spesso la replicazione simmetrica



24

Trigger di replicazione

- Catturano le variazioni ai dati nelle tabelle DELTA-PLUS E DELTA-MINUS in modo trasparente alle applicazioni
- Esempio:

```
CREATE TRIGGER CAPTURE-INS  
AFTER INSERT ON PRIMARY  
FOR EACH ROW  
INSERT INTO DELTA-PLUS VALUES (NEW.*)
```

```
CREATE TRIGGER CAPTURE-DEL  
AFTER DELETE ON PRIMARY  
FOR EACH ROW  
INSERT INTO DELTA-MINUS VALUES (OLD.*)
```

```
CREATE TRIGGER CAPTURE-UPD  
AFTER UPDATE ON PRIMARY  
FOR EACH ROW  
BEGIN  
INSERT INTO DELTA-PLUS VALUES (NEW.*)  
INSERT INTO DELTA-MINUS VALUES (OLD.*)  
END
```