

Parte 3

PUNTO A. Progetto fisico e sua validazione

a) Descrizione in linguaggio naturale e il codice SQL sviluppato per implementare le interrogazioni del carico di lavoro

1° QUERY: seleziona tutte le rilevazioni effettuate in data '2003\06\22'

```
SELECT *  
FROM Rilevazioni  
WHERE timeRil = '2003\06\22';
```

2° QUERY: seleziona il nome comune delle specie per le quali sono state piantate delle repliche nel gruppo n° 20.

```
SELECT NComune  
FROM Piante JOIN Specie ON Piante.Nscientifico = Specie.Nscientifico  
WHERE Piante.gruppo = 20;
```

3° QUERY: seleziona il nome degli orti “nel pulito” in cui sono presenti 10 sensori, e che hanno una superficie maggiore di 100 m².

```
SELECT Nome  
FROM Orti  
WHERE Nsensori = 10 AND Superficie > 100 AND Pulito = true;
```

b) Progetto fisico, contenente l'elenco degli indici che si intendono creare per le interrogazioni contenute nel carico di lavoro (specificando relazione di riferimento e chiave di ricerca), il loro tipo (ordinato/hash, clusterizzato/non clusterizzato) e la motivazione che ha portato alla loro creazione.

1° QUERY: abbiamo deciso di utilizzare un indice hash sulla relazione “Rilevazioni”, ed in particolare sull'attributo “timeRil”. La relazione potrebbe ipoteticamente essere la più popolata sulla base dati essendo che per ogni pianta potrebbero essere effettuate più rilevazioni in giorni ed ore diverse, ma al contempo nello stesso giorno potrebbero essere effettuate molte rilevazioni.

Un indice hash risulta essere il più appropriato, difatti avendo una corrispondenza diretta tra chiavi e valori siamo in grado di accedere in modo efficiente a tutte le rilevazioni effettuate in un certa data, questo senza dover memorizzare fisicamente l'indice.

Si è scelto di non clusterizzare il file dei dati su tale indice in quanto, come detto la relazione è molto dinamica e potrebbe richiedere numerosi inserimenti.

La clusterizzazione tramite indice hash potrebbe richiedere una ridistribuzione frequente dei dati per mantenere l'equilibrio della distribuzione hash. Ciò potrebbe comportare un carico significativo sul sistema durante gli aggiornamenti e potenzialmente influire sulle prestazioni.

CREATE INDEX idxRil ON Rilevazioni USING HASH (timeRil);

2° QUERY: per ottimizzare l'operazione di join tra relazioni "Piante" e "Specie" abbiamo provato ad introdurre un indice ordinato clusterizzato e primario su "Specie", sull'attributo "Nscientifico".

Questo al fine di ottenere l'utilizzo dell'index nested loop join, abbiamo però riscontrato che il sistema continuava a prediligere l'utilizzo di hash inner join, probabilmente a cause delle grosse dimensioni delle tabelle coinvolte.

Abbiamo quindi scelto di utilizzare un indice ordinato clusterizzato sull'attributo "gruppo" di "Piante", in modo da ottenere un'ottimizzazione sulla ricerca delle piante appartenenti ad un dato gruppo.

La clusterizzazione del file dati rispetto a tale attributo rende praticamente immediata la ricerca delle piante associate a determinati gruppi.

CREATE INDEX idxPiante ON Piante(gruppo);
CLUSTER Piante USING idxPiante;

3° QUERY: per la relazione "Orti" è stato scelto l'indice ordinato non clusterizzato su "Nsensori" e "Superficie".

L'utilizzo di un indice multiattributo ci permette di andare ad effettuare la ricerca per uguaglianza rispetto al numero di sensori in modo efficiente, e da qui di ricercare per intervallo le tuple sul secondo attributo.

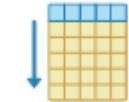
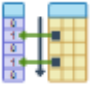

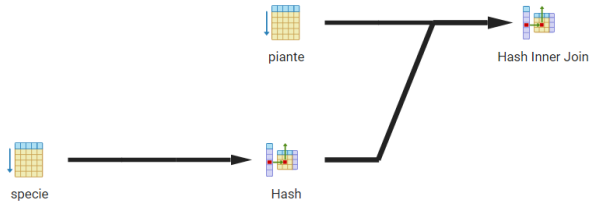
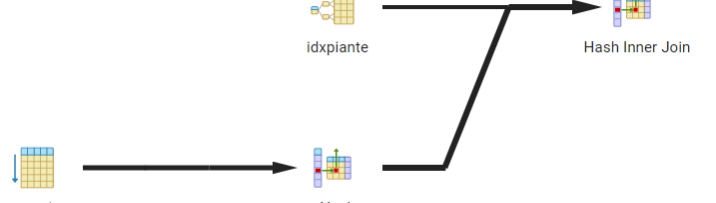
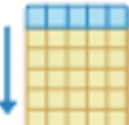

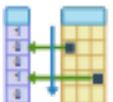
CREATE INDEX idxOrti ON Orti (Nsensori, Superficie);

c) Una tabella che, per ogni tabella coinvolta nelle operazioni del carico di lavoro, riporti il numero di tuple inserite e la dimensione in blocchi della tabella.

Relazioni	Numero Tuple	Numero Pagine
Orti	770	19
Specie	830	19
Piante	666	20
Rilevazioni	1000	20

- d) Descrizione dei piani di esecuzione scelti dal sistema (prima e dopo la creazione dello schema fisico) per le interrogazioni contenute nel carico di lavoro, riportando la visualizzazione del piano di esecuzione prodotta da PostgreSQL e i tempi di esecuzione (prima e dopo la creazione dello schema fisico); confrontare i piani e i tempi ottenuti nei due casi, fornendo una giustificazione per i risultati ottenuti.

(I piani di esecuzione riportati in tabella seguono l'ordine delle query per il carico di lavoro)

SENZA INDICE	CON INDICE
 <p>Viene effettuata una scansione sequenziale con filtro su 'timeRil'</p> <p>Tempo di accesso: 0,116 ms</p> <p>rilevazioni</p>	 <p>idxril</p>  <p>rilevazioni</p> <p>Utilizzando un indice hash su 'timeRil' si ha un costo di 0,017 ms dato dal tempo impiegato per scansionare l'indice bitmap e individuare gli identificatori dei blocchi che soddisfano la condizione specifica. In ulteriori 0,009 ms scansiona le pagine di memoria individuate nel punto precedente, si ha un tempo totale di accesso pari a 0,026 ms.</p>
 <p>Costo totale 0,389 ms, di cui 0,133 ms sono stati impiegati per effettuare la scansione di piante.</p>	 <p>Introducendo l'indice su Piante, si semplifica notevolmente l'accesso con filter su "gruppo" per tale relazione, che ha un costo di soli 0,013 ms, mentre il costo totale per l'operazione si riduce a 0,266 ms</p>
 <p>Viene effettuata una scansione sequenziale su orti in 0,151ms.</p> <p>orti</p>	 <p>idxorti</p>  <p>orti</p> <p>Utilizzando un indice ordinato su (Nsensori, Superficie) viene effettuato un Index Scan sul primo attributo per uguaglianza, e sul secondo per intervallo. Si ha un tempo di esecuzione di 0,032 ms</p>

PUNTO B. Descrizione della politica di controllo dell'accesso scelta

Sono stati definiti i ruoli per **insegnante**, **gestore globale del progetto**, **referente di scuola** e **studente**. Avendo disambiguato la specifica tenendo conto delle scuole come istituti in sé, il referente di istituto viene incluso nel referente di scuola.

Segue la descrizione della politica di controllo dell'accesso scelta.

Il **gestore globale del progetto** è colui che necessita di tutti i privilegi su ogni relazione, deve poter modificare secondo le sue necessità la base dati, poter assegnare privilegi ad altri utenti o ad altri ruoli se necessario.

Rappresentando il ruolo con più privilegio, il gestore del progetto eredita il ruolo immediatamente sottostante ovvero quello del rappresentante della scuola.

Rispetto al **rappresentante della scuola** abbiamo deciso di mantenere ancora un certo numero di privilegi in quanto per come abbiamo interpretato le specifiche, sarà proprio la scuola ad occuparsi della gestione di chi partecipa al progetto, quindi delle relative classi o persone, degli orti che possiede, dei dispositivi presenti in ogni orto ecc.

L'unico privilegio rispetto al quale abbiamo imposto una limitazione è rispetto alle altre scuole, un rappresentante potrà dunque soltanto visualizzare dati di altre scuole senza ovviamente modificarli, inserirli o eliminarli.

Il rappresentante eredita il ruolo di insegnante.

L'**insegnante** a sua volta avrà un numero di privilegi ridotto, non ha poteri di modifica inserimento o cancellazione su orti, classi, persone, scuole o dispositivi, in quanto risultano essere tutti campi al di fuori del suo dominio, nei quali non deve poter metter mano se non per visualizzare eventualmente dati utili ai fini del progetto.

Eredita il ruolo di studente.

Lo **studente** sarà l'individuo con il minor numero di privilegi, in particolare non ha nessun privilegio rispetto a persone o scuole, rappresentano informazioni al di fuori della sua competenza.

Dovrà essere in grado di inserire rilevazioni, e memorizzare eventualmente quando ha effettuato l'inserimento, e che informazioni ha inserito.

Allo stesso modo garantiamo la possibilità di inserire nella base dati eventuali specie introdotte, e delle repliche per tali specie.

Limitiamo totalmente il privilegio di delete allo studente, sarà compito dei ruoli superiori gestire questo aspetto per motivi di sicurezza. L'unica relazione sulla quale può effettuare una delete è piante, saranno potenzialmente gli studenti ad avere il pieno controllo sulle repliche che piantano.

Schema dei privilegi rispetto agli utenti

	Nomeutente1 (gestoreGlobaleProgetto)	Nomeutente2 (referenteScuola)	Nomeutente3 (insegnante)	Nomeutente4 (studente)
Scuole	SELECT INSERT DELETE UPDATE	SELECT	SELECT	\
Persone	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT	\
Classi	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT	SELECT
Orti	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT	SELECT
Specie	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT INSERT UPDATE
Gruppi	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT
Piante	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE
Dispositivi	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT	SELECT
Rilevazioni	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT INSERT
Inserisce	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT INSERT UPDATE
EffettuaRil	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT INSERT UPDATE
Informazioni	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT INSERT DELETE UPDATE	SELECT INSERT UPDATE