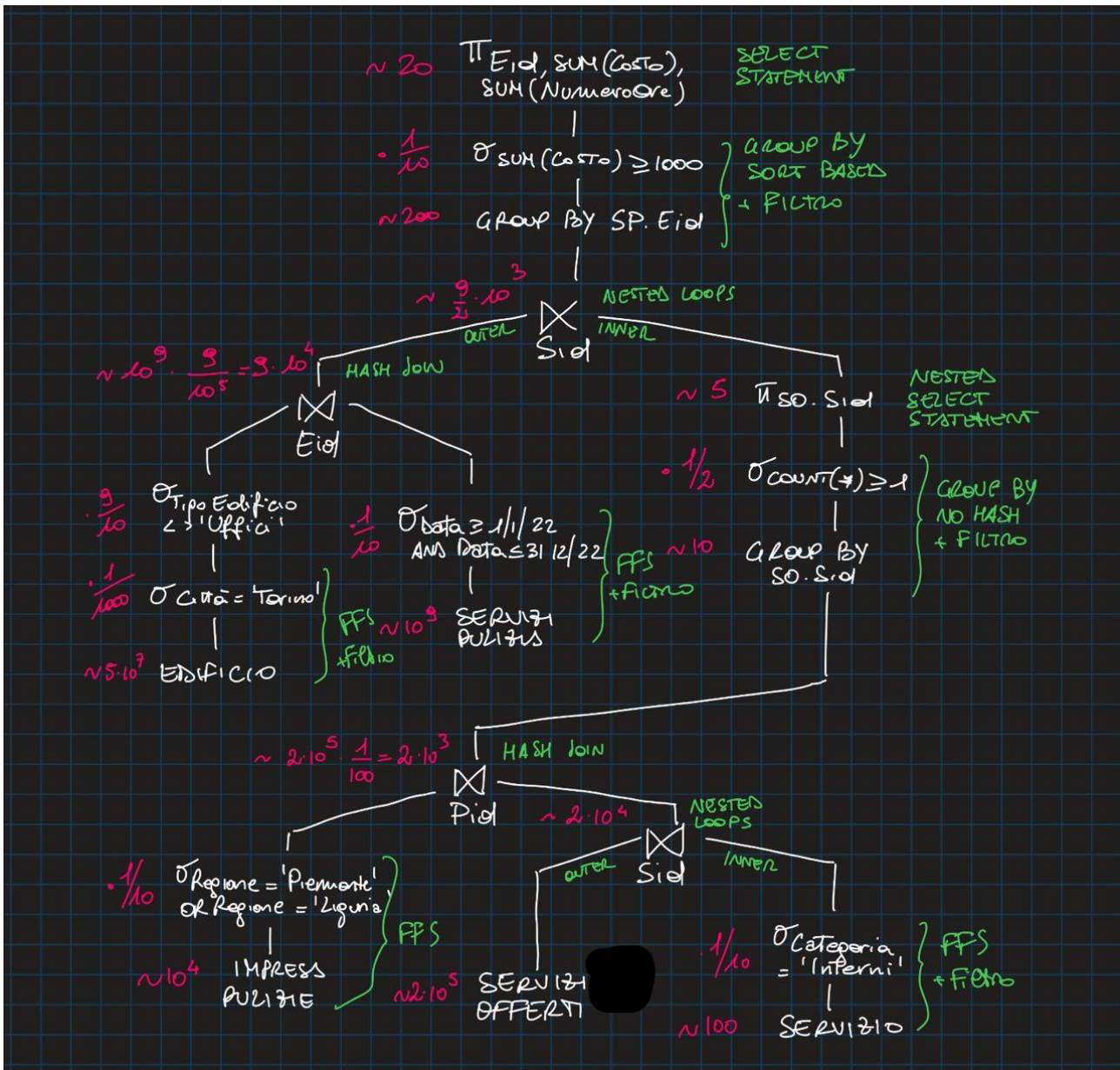


Quaderno 3

Punto 1

Si scriva l'espressione algebrica corrispondente, indicando le operazioni svolte, la cardinalità e la selettività di ogni operazione. Dove necessario, si ipotizzi la distribuzione dei dati. Discutere la possibilità di anticipare l'operatore GROUP BY.



Si è fatta assunzione di uniformità delle distribuzioni sui seguenti attributi: SERVIZIO.Categoria, IMPRESA_PULIZIE.Regione, EDIFICIO.Città, EDIFICIO.TipoEdificio, SERVIZI_PULIZIA.Data.

Per quanto riguarda l'anticipo della group by:

- La group by della query interna non può essere anticipata a meno di cabiarne gli attributi e quindi calcolandola su Pid e Sid. Tuttavia, la sua utilità sarebbe nulla dal momento che

Ovunque la si metta, il fattore di riduzione sarebbe 1 a causa della tabella SERVIZI_OFFERTI in cui ogni riga rappresenta univocamente Pid e Sid.

- La group by della query esterna non può essere anticipata a meno di modificarne la firma poiché si trova subito dopo l'esecuzione del semi join sulla chiave Pid, che essa non contempla.

Nota: con FFS si intende full fast scan. Anche se non richiesto, dal primo punto, è stato scritto un piano di piano di esecuzione senza l'uso di indici come base di partenza.

Punto 2

Si scelgano le strutture fisiche accessorie per migliorare le prestazioni dell'interrogazione. Si motivi la scelta e si definisca il piano di esecuzione (ordine e tipo dei join, accesso alle tabelle e/o indici, etc.).

Strutture Accessorie

IMPRESA-PULIZIE:

- Indice su Regione
 - Selettività: buona, 1/10
 - Dimensioni tabella: grande, ca 1000
 - ⇒ Un indice è sensato
 - Aiuto operazioni successive: no oltre alla select
 - Regione non è chiave primaria, cardinalità bassa 20 (bitmap più efficiente di b+-tree se la cardinalità dell'attributo è minore di 32)
 - La selezione è di tipo uguaglianza ma è complessa (doppio predicato con OR)
 - ⇒ Indice secondario unclustered non coprente su Regione di tipo bitmap

```
CREATE BITMAP INDEX idx_regione_bitmap
ON IMPRESA_PULIZIE(Regione);
```

SERVIZI-OFFERTI

- Indice su Sid
 - Selettività: nessuna selezione immediata
 - Dimensioni tabella: grande 2×10^5
 - Aiuto operazioni successive: No
 - ⇒ Non conviene fare un indice

SERVIZIO

- Indice su Sid
 - Selettività: buona, 1/10
 - Dimensioni tabella: piccola
 - ⇒ Non conviene fare una indice

EDIFICIO

- Indice su Città
 - Selettività: Molto buona, 1/1000
 - Dimensioni tabella: Molto grande 5×10^7
 - ⇒ Conviene fare un indice
 - Aiuta operazioni successive: no
 - Cardinalità attributo: 1000 => bitmap non sensato
 - ⇒ Indice secondario unclustered non coprente su Città di tipo Hash

```
CREATE INDEX idx_citta_hash  
ON EDIFICIO(Città)  
USING HASH;
```

- Indice su TipoEdificio
 - Selettività: molto bassa, 9/10
 - ⇒ Non conviene fare un indice

SERVIZI-PULIZIA

- Indice su Data
 - Selettività: buona, 1/10
 - Dimensioni tabella: Molto grande, 10^9
 - ⇒ Un indice ha senso
 - Il predicato non è di uguaglianza esatta
 - La data viene usata solo nella selezione immediatamente successiva
 - ⇒ Indice secondario unclustered non coprente su Data di tipo B+-tree

```
CREATE INDEX idx_data_btree  
ON SERVIZI_PULIZIA(Data);
```

Piano di esecuzione finale

