

Relazione Homework 1 Tecnologie e Architetture per la Gestione dei Dati

Inizialmente sono state create due relazioni R1 e R2 descritte come segue:

create table r1 (a numeric not null primary key, b numeric, c numeric);

create table r2 (d numeric not null primary key, e numeric, f numeric);

Sono state inserite in maniera randomica 1000000 e 4000000 di ennuple rispettivamente in questo modo:

```
Query Query History
24 delete from r1;
25 create or replace function loadr1()
26 returns void as $$
27 begin
28 for i in 1..1000000 loop
29 insert into r1 (a,b,c)
30 values(i, floor(random()*300000+1),
31 floor(random()*400000+1);
32 end loop;
33 end;
34 $$ language plpgsql;
35 select loadr1();
36
```

Data Output Messages Notifications

	loadr1	void
1		

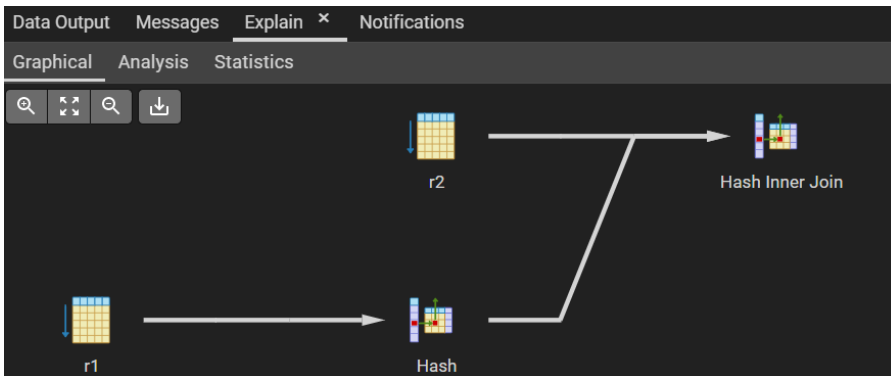
```
Query Query History
12 create or replace function loadr2()
13 returns void as $$
14 begin
15 for i in 1..4000000 loop
16 insert into r2 (d,e,f)
17 values(i, floor(random()*100+1),
18 floor(random()*1000+1);
19 end loop;
20 end;
21 $$ language plpgsql;
22 select loadr2();
23
24
```

Data Output Messages Notifications

	loadr2	void
1		

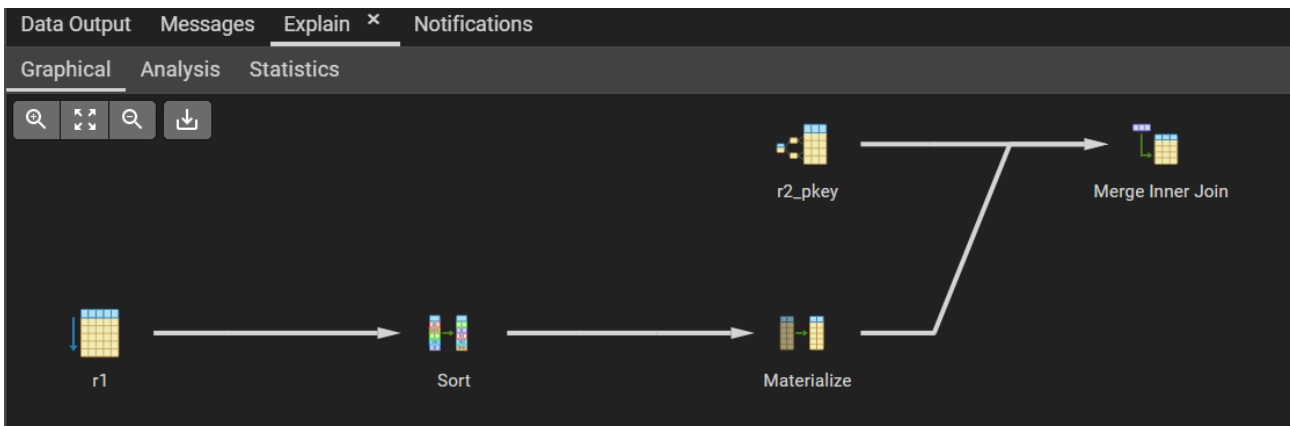
È stato fatto un join tra la relazione R1 e la relazione R2 con la condizione $C=D$.

Il risultato viene mostrato in figura.

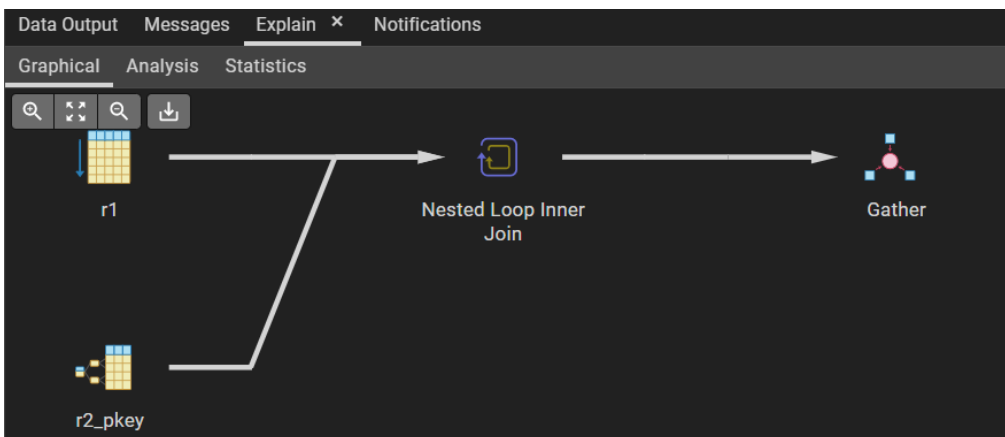


In questo caso viene scelto di fare l'hash sulla relazione R1 in quanto è la relazione con dimensioni minori. Successivamente viene eseguito il join con la relazione R2.

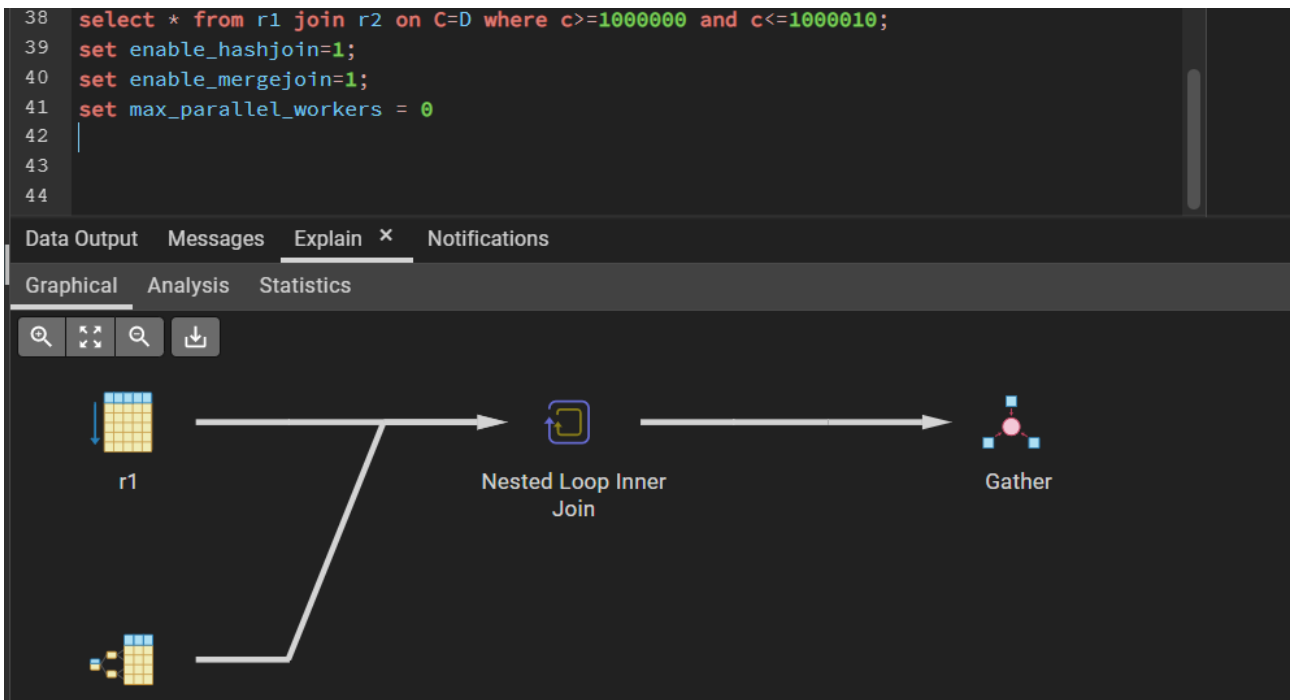
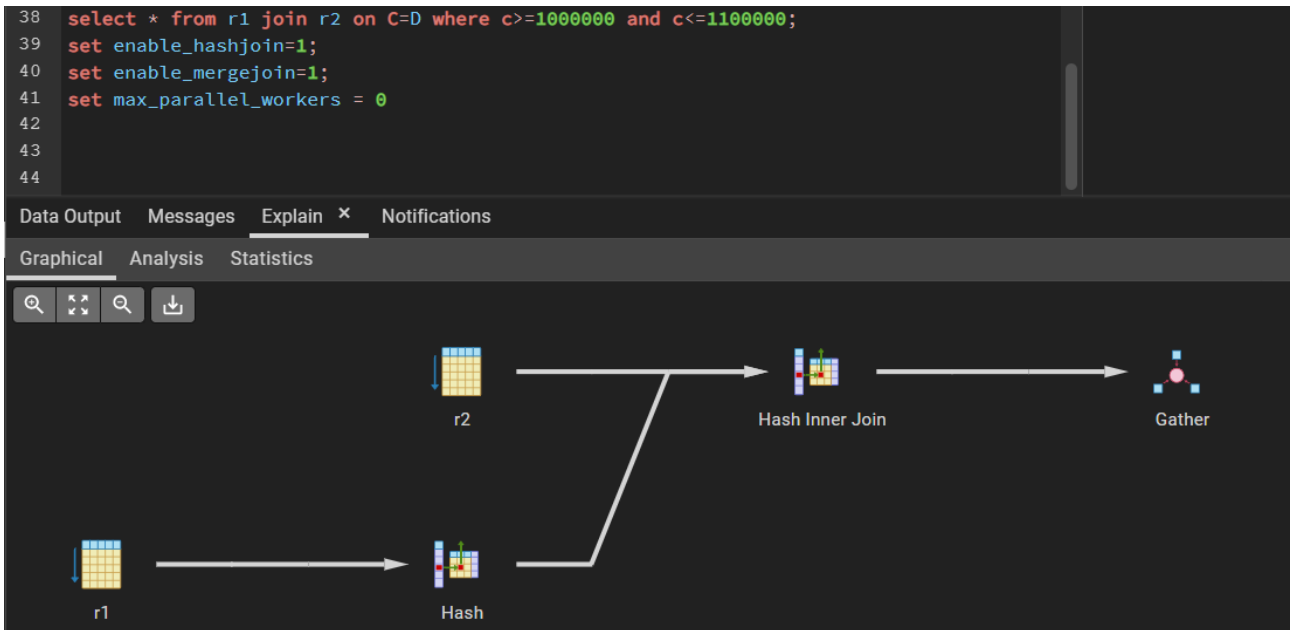
Disabilitando l'hash join si può evincere la soluzione adottata, come mostrato in figura.



Si può disabilitare anche il merge join per vedere nuove soluzioni scelte.



Ulteriori test si possono fare inserendo una condizione di where, facendo una selezione. In particolare si può notare che variando di 100000 ennuple il valore di c, il DBMS passa da una scelta di hash join a una scelta di nested loop.



Infine si può aggiungere un indice sull'attributo C, in cui il tempo di esecuzione della query diminuisce notevolmente.

```
38 select * from r1 join r2 on C=D where c>=1000000 and c<=1000010;
39 set enable_hashjoin=1;
40 set enable_mergejoin=1;
41 set max_parallel_workers = 0
42
43 create index i12 on r1(C);
44
```

Data Output Messages Explain **x** Notifications

Graphical Analysis Statistics

The diagram illustrates the execution plan for the query. It shows two input nodes on the left: 'i12' (represented by a yellow and blue grid icon) and 'r2_pkey' (represented by a yellow and blue grid icon). Arrows from both nodes converge into a single arrow pointing to a 'Nested Loop Inner Join' node on the right, which is represented by a blue square icon with a white arrow. The interface includes a toolbar with search, zoom, and download icons.