Clauza Connect By, Ierarhii, Cicluri, Recursivitate:

Referat Sisteme de gestiune a bazelor de date

Tender Laura - Maria

Grupa 234

December 1, 2020

Introducere

În acest referat vom descrie o metodă de abordare a cererilor ierarhice cu ajutorul clauzei *CONNECT BY*. Vom analiza parametrii care pot fi folosiți împreună cu această clauză pentru a afla mai multe informații despre ierarhie. Vom discuta modul în care se comportă ciclurile și funcțiile recursive infinite.

Descrierea datelor și a problemei abordate

Exemplele prezentate în acest referat folosesc tabelul *Persoana* care a fost creat folosind următoarele comenzi:

ADD CONSTRAINT UQ_CNP UNIQUE (CNP);

În tabel au fost inserate date astfel:

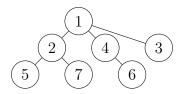
COMMIT;

Datele din tabel sunt următoarele:

				∜ NUME			
1	1	(null)	1410721106449	Popa	Ioan	0740000000	Buzau
2	2	1	2651210226968	Popa	Maria	0741111111	Iasi
3	3	1	1680217228786	Ionescu	Marcel	074222222	Iasi
4	4	1	1710428229854	Popa	George	0743333333	Iasi
5	5	2	1910815445766	Ionescu	Andrei	0750000000	Bucuresti
6	6	4	5011003127121	Popa	Alexandru	0760000000	Cluj
7	7	2	6011110417243	Ionescu	Ioana	0751111111	Bucuresti

Între persoanele din tabel există o relație părinte - copil. Astfel persoanele din tabel pot fi reprezentate printr-un arbore. Acest arbore ne va ajuta să observăm mai bine modul

în care funcționeaza cererile din cadrul acestui referat.



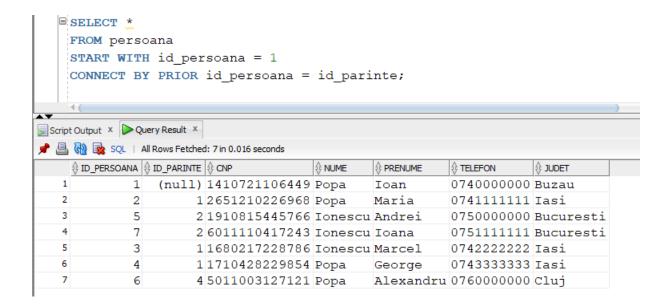
Problema generală pe care o vom aborda în exemplele următoare este de a selecta persoanele din baza de date pe baza ierarhiei părinte - copil.

Introducere în CONNECT BY

Fie următorul cod care obține următorul rezultat: SELECT * FROM persoana CONNECT BY PRIOR id persoana = id parinte; Script Output × Query Result × 📌 搗 🙌 퀋 SQL | All Rows Fetched: 16 in 0.017 seconds ⊕ NUME JUDET 1 12651210226968 Popa Maria 0741111111 Iasi 2 5 2 1910815445766 Ionescu Andrei 0750000000 Bucuresti 7 3 2 6011110417243 Ionescu Ioana 07511111111 Bucuresti 3 11680217228786 Ionescu Marcel 0742222222 Iasi 5 4 1 1710428229854 Popa George 0743333333 Iasi 6 6 4 5011003127121 Popa Alexandru 0760000000 Cluj 7 2 1910815445766 Ionescu Andrei 07500000000 Bucuresti 7 8 2 6011110417243 Ionescu Ioana 07511111111 Bucuresti 6 9 4 5011003127121 Popa Alexandru 0760000000 Cluj 1 10 (null) 1410721106449 Popa Ioan 0740000000 Buzau 11 2 12651210226968 Popa 0741111111 Iasi Maria 5 12 2 1910815445766 Ionescu Andrei 07500000000 Bucuresti 13 7 2 6011110417243 Ionescu Ioana 07511111111 Bucuresti 14 3 0742222222 Iasi 11680217228786 Ionescu Marcel 15 4 1 1710428229854 Popa George 0743333333 Iasi 16 4 5011003127121 Popa Alexandru 0760000000 Cluj

O primă observație este aceea că apar duplicate, obținem mai multe rezultate decât numărul de rânduri din tabelul *Persoana*. Vom înțelege mai bine de ce obținem acest rezultat în următoarele exemple.

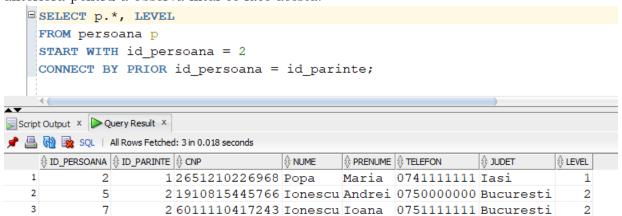
Ne dorim să nu obținem duplicate și să obținem rezultatele pornind din rădăcină. Astfel vom folosi START WITH.



Folosind START WITH obținem toate nodurile subarborelui o singură dată, în ordinea dată de parcurgere.

LEVEL

Folosind LEVEL putem observa mai bine cum functioneaza CONNECT BY, PRIOR si ce se intampla atunci cand nu folosim START WITH. Vom adauga LEVEL la comanda anterioră pentru a observa întâi ce face acesta.



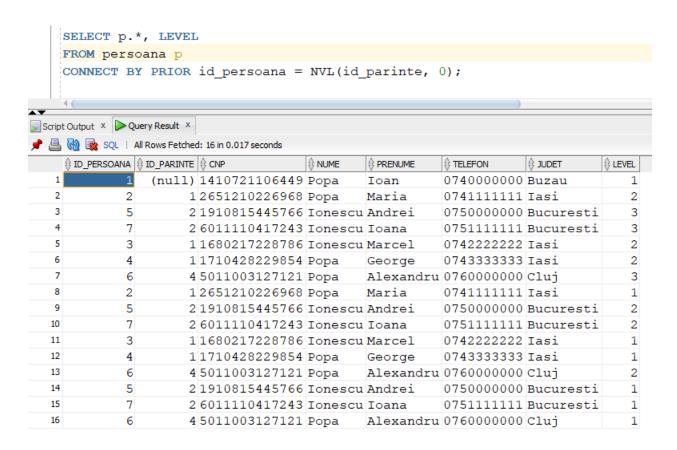
Obținem subarborele rădăcinii precizate cu ajutorul START WITH, iar LEVEL indică nivelul nodului în subgraf, rădăcina primește nivelul 1, fiii acesteia nivelul 2 și tot așa.

Vom folosi LEVEL pentru cererea fără START WITH și vom obține următorul rezultat.

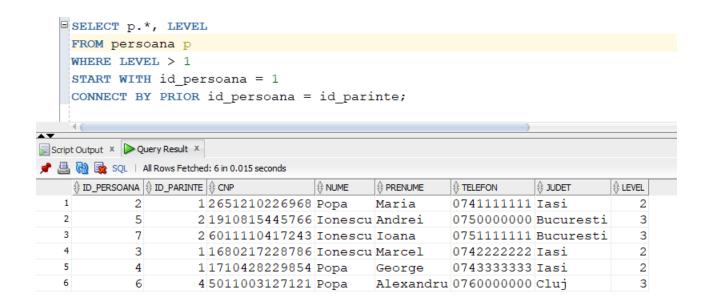
S	ELECT p.*,	, LEVEL						
FI	ROM persoa	ana p						
	CONNECT BY PRIOR id persoana = id parinte;							
		-						
· '								
Script O	utput × Quer	ry Result X						
P 🖺 🛭	🤰 🌉 SQL AllI	Rows Fetched: 16 in 0.031 seconds						
\$	ID_PERSOANA ():	ID_PARINTE & CNP	∜ NUME	♦ PRENUME	∯ TELEFON			
1	2	1 2651210226968	Popa	Maria	0741111111	Iasi	1	
2	5	2 1910815445766	Ionescu	Andrei	0750000000	Bucuresti	2	
3	7	2 6011110417243	Ionescu	Ioana	0751111111	Bucuresti	2	
4	3	1 1680217228786	Ionescu	Marcel	074222222	Iasi	1	
5	4	11710428229854	Popa	George	0743333333	Iasi	1	
6	6	4 5011003127121	Popa	Alexandru	0760000000	Cluj	2	
7	5	2 1910815445766	Ionescu	Andrei	0750000000	Bucuresti	1	
8	7	2 6011110417243	Ionescu	Ioana	0751111111	Bucuresti	1	
9	6	4 5011003127121	Popa	Alexandru	0760000000	Cluj	1	
10	1	(null) 1410721106449	Popa	Ioan	0740000000	Buzau	1	
11	2	1 2651210226968	Popa	Maria	0741111111	Iasi	2	
12	5	2 1910815445766	Ionescu	Andrei	0750000000	Bucuresti	3	
13	7	2 6011110417243	Ionescu	Ioana	0751111111	Bucuresti	3	
14	3	1 1680217228786	Ionescu	Marcel	0742222222	Iasi	2	
15	4	1 1710428229854	Popa	George	0743333333	Iasi	2	
16	6	4 5011003127121	Popa	Alexandru	0760000000	Cluj	3	

Parcurgerea porneste din primul rând din tabel care îndeplinește condiția (1 nu indeplineste condiția întrucât are id parinte NULL). În parcurgere, nivelul celorlalte noduri se raporteaza la cele vizitate deja, fiii primind nivelul parintelui + 1. După ce parcurgerea din nodul curent a luat sfârșit, se pornește o nouă parcurgere de la următorul nod din tabel. Din acest motiv 1, 2, 3 apar cu nivelul 1 în rezultat și obținem duplicate. Observăm că parcurgerea pornită din 1 nu obține rezultate contradictorii.

Vom face ca prima parcurgere să pornească cu primul rând din tabel cu ajutorul NVL.



Următoarea problemă - Obțineți strămoșii persoanei cu id-ul 1 poate fi rezolvată ușor folosind LEVEL în clauza WHERE.

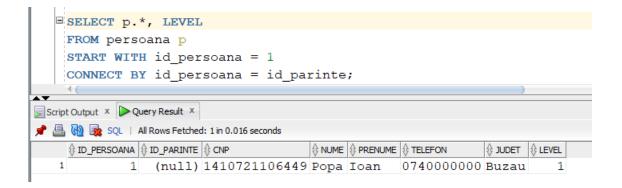


PRIOR

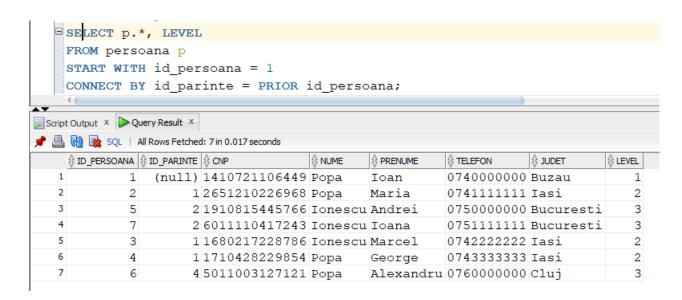
În cererile anterioare am folosit PRIOR însă nu discutat ce face acesta și cum îl putem folosi. Pentru început, vom observa ce se întâmplă în cereri ce nu conțin PRIOR.

	absenta	a prior								
	SELECT p.*, LEVEL									
	FROM persoana p									
	CONNECT BY id persoana = id parinte;									
A.	1									
	Script Output × Query Result ×									
_	≠ 🔠 🙀 🔯 SQL All Rows Fetched: 7 in 0.016 seconds									
		D_PARINTE	∯ CNP		∜ NUME				\$ LEVEL	
1	1	(null)	141072	1106449	Popa	Ioan	0740000000	Buzau	1	
2	2	1	265121	.0226968	Popa	Maria	0741111111	Iasi	1	
3	3	1	168021	.7228786	Ionescu	Marcel	074222222	Iasi	1	
4	4	1	171042	8229854	Popa	George	0743333333	Iasi	1	
5	5	2	191081	5445766	Ionescu	Andrei	0750000000	Bucuresti	1	
6	6	4	501100	3127121	Popa	Alexandru	0760000000	Cluj	1	
7	7	2	601111	.0417243	Ionescu	Ioana	0751111111	Bucuresti	1	

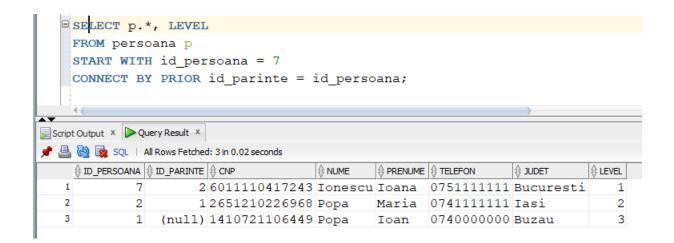
Observăm că obținem fiecare nod cu nivelul 1, absența *PRIOR* face ca cererea să nu se mai raporteze la nodul anterior, deci nu se mai întâmplă o parcurgere. Dacă indicăm din ce nod dorim să înceapă parcurgerea obținem doar acel nod.



Putem așeza PRIOR și în partea dreaptă a clauzei CONNECT BY.



Penrtru a obține o parcurgere ascendentă, de la un copil către părinte putem scrie o astfel de cerere:



Cicluri

Dacă graful nostru ar conține un ciclu, parcurgerea ar putea merge la infinit sau să își piardă însemnătatea. Am adaugat în mod eronat un ciclu în tabel pentru a vedea ce rezultate obținem.

UPDATE Persoana

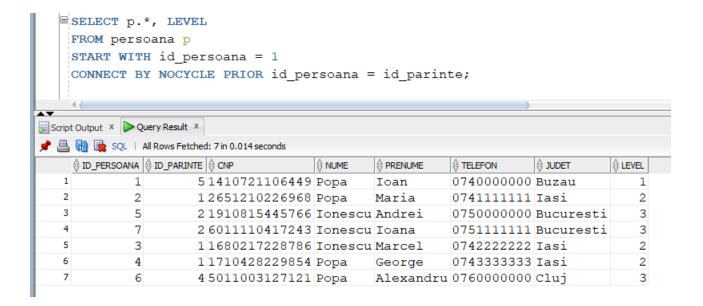
```
SET id_parinte = 5
WHERE id_persoana = 1;
COMMIT;
```

Rulând aceeași cerere de pănă acum în acest caz obținem eroarea CONNECT BY loop in user data. Deci algoritmul detectează existența unui ciclu și aruncă o erorare în acest caz.

```
SELECT p.*, LEVEL
FROM persoana p
START WITH id_persoana = 1
CONNECT BY PRIOR id_persoana = id_parinte;

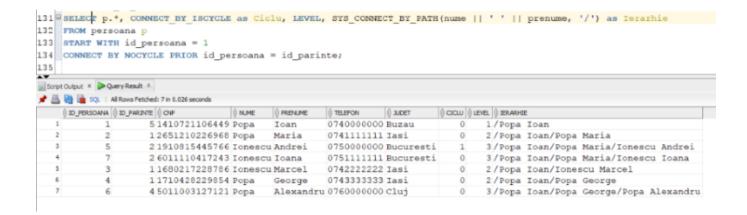
Script Output x Query Result x
Script Output x Query Result x
Script Output x SQL | Executing: SELECT p.*, LEVELFROM persoana p START WITH id_persoana = 10
ORA-01436: CONNECT BY loop in user data
01436. 00000 - "CONNECT BY loop in user data"
*Cause:
*Action:
```

Dacă dorim să ignorăm ciclul și să obținem rezultatele putem folosi atributul NOCYCLE.



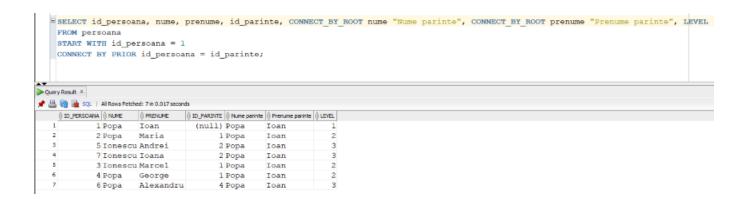
Pentru a determina nodurile care închid un ciclu putem folosi atributul CONNECT BY IS-CYCLE. În plus, pentru a afișa informații despre parcurgere putem folosi SYS CONNECT

BY PATH.

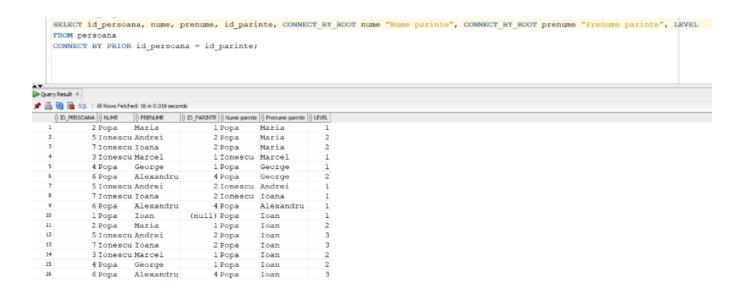


CONNECT BY ROOT, CONNECT BY ISLEAF

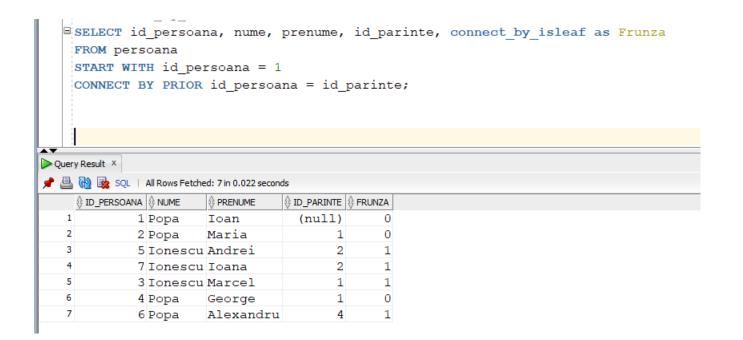
 $CONNECT\ BY\ ROOT\$ ajută la obținerea de informație despre rădăcină, nodul din care a pornit parcurgerea.



Observăm că pentru fiecare nod obținem informația despre nodul din START WITH.



Pentru a afla dacă un nod este frunză putem folosi CONNECT BY ISLEAF. Astfel aflăm persoanele fără copii.



Parcurgerea

Vom descrie pașii pe care Oracle îi urmează pentru a crea ierarhia folosind CONNECT

Sunt selectate rândurile care îndeplinescu condiția din START WITH. Apoi Oracle selectează rândurile copil pentru fiecare rădăcină. Copiii trebuie să îndeplinească condiția din CONNECT BY față de o rădăcină. Dacă cererea conține o clauza WHERE atunci Oracle evalueaza fiecare rând și le păstrează pe cele care îndeplinesc condiția. 1

Observăm că rezultatul este o parcurgere Depth First Search (DFS).

Recursivitate

Fie functia recursiva factorial din laboratorul 4 PL/SQL (8).

```
RETURN INTEGER IS

BEGIN

IF (n=0) THEN RETURN 1;

ELSE RETURN n*factorial_lte(n-1);

END IF;

END factorial_lte;

/

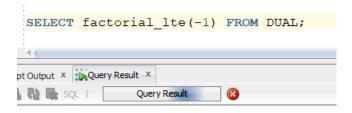
SELECT factorial_lte(5) FROM DUAL;

ript Output x Query Result x

FACTORIAL_LTE(5)

1 120
```

Am apelat funcția factorial pentru o valoare negativă.



^{1.} Hierarchical Queries, Database SQL Reference, Oracle Documentation

Dacă subprogramul nu ajunge la o condiție terminală, recursia continuă până PL/SQL rămâne fără memorie și aruncă excepția predefinită STORAGE ERROR. 2

Putem rescrie funcția factorial adaugând o excepție pentru numere negative.

```
173 CREATE OR REPLACE FUNCTION factorial_lte(n NUMBER)
174
           RETURN INTEGER IS
175
               negative_number EXCEPTION;
176
               PRAGMA EXCEPTION_INIT (negative_number, -20001);
177
          BEGIN
178
               IF (n < 0) THEN
179
                    raise_application_error(-20001,'Nu poate fi calculat factorialul unui numar negativ');
180
               ELSE
181
                    IF (n=0) THEN RETURN 1;
182
                    ELSE RETURN n*factorial_lte(n-1);
183
                    END IF;
184
               END IF;
185 END factorial_lte;
186 /
187
188
     SELECT factorial lte(-1) FROM DUAL;
Script Output × DQuery Result ×
📌 📇 🍓 🅦 SQL | Executing:SELECT factorial_lte(-1) FROM DUAL in 0 seconds
ORA-20001: Nu poate fi calculat factorialul unui numar negativ
ORA-06512: at "GRUPA234.FACTORIAL_LTE", line 7
```

Bibliografie

Hierarchical Queries in Oracle, Oracle-Base

Connect By, Snowflake Documentation

CONNECT BY clause in Oracle, Oracle

PL/SQL Raise Exceptions, Oracle Tutorial

^{2.} Using Recursive PL/SQL Subprograms, Database PL/SQL Language Reference, Oracle Documentation