



Indexes

- Definiție
- Organizarea stocării informației
- B-Tree
- Tipuri de indecși
- Crearea și ștergerea index

Index - definiție:

Indexul este o structură de date asociată cu un tabel, care permite accesul rapid la informații și îmbunătățește performanța interogărilor bazei de date, ajutând DBMS-ul să găsească rapid rânduri potrivite criteriilor de căutare (analogic cuprinsului unui cărți).

În lipsa indexului, căutarea se efectuează prin scanarea întregului tabel, rând cu rând, ceea ce poate fi foarte ineficient și consumator de timp, mai ales pentru tabele mari.

Totodată, mulți indecși pot încetini operațiunile de inserare, actualizare și ștergere, deoarece fiecare modificare a datelor necesită actualizarea tuturor indecșilor asociați cu tabelul, ceea ce poate afecta performanța.

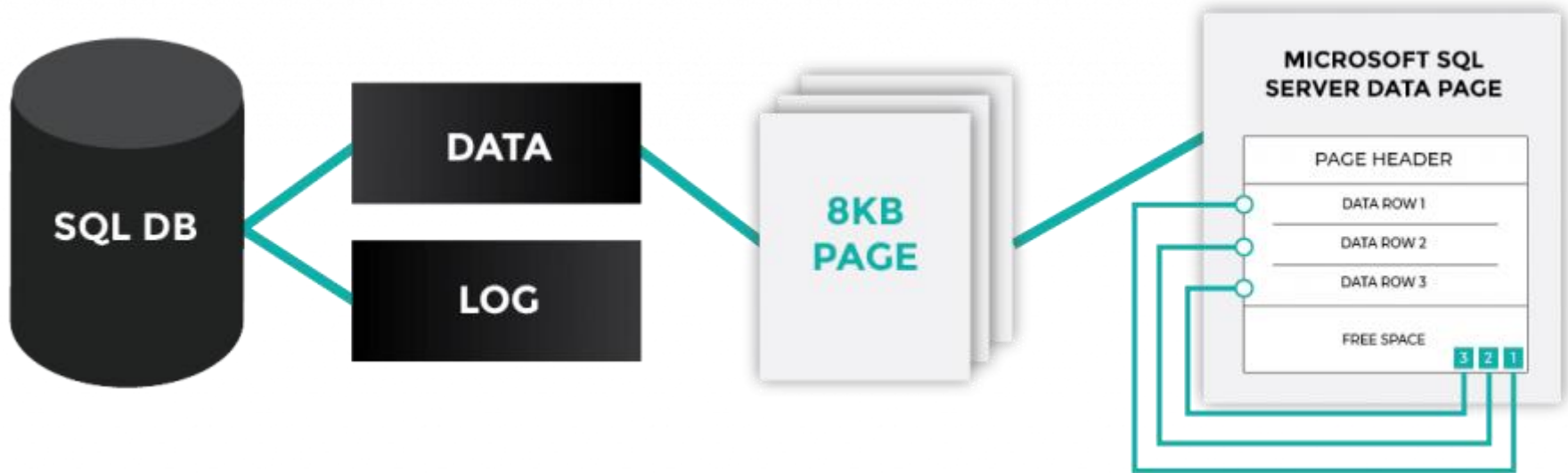
Indecșii trebuie de creat cu prudență, ținând cont de următoarele aspecte:

- indecșii sunt generați automat pentru câmpurile definite ca **chei primare**;
- indecșii trebuie de creat pentru câmpurile:
 - a) care sunt frecvent utilizate în căutări;
 - b) cu un număr minim de valori duplicate și NOT NULL.

Organizarea Stocării Informațiilor în SQL Server:

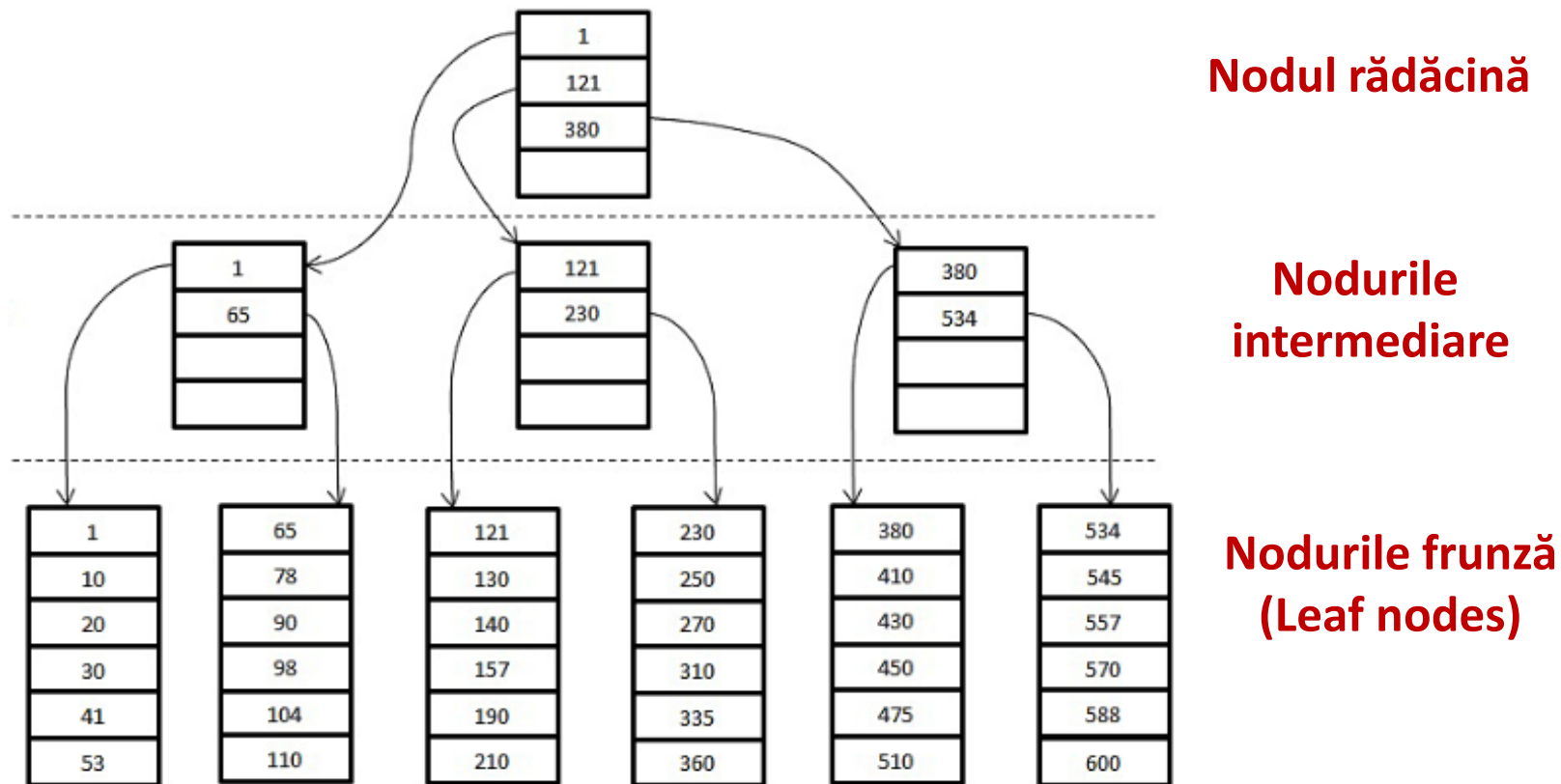
Înainte de a descrie structura indexului, este important să înțelegem cum sunt stocate datele în baza de date. Unitatea minimă de alocare este **pagina**, fiecare având 8 KB.

Există două tipuri de pagini - paginile de date și paginile de index.



B-Tree (Balanced Tree):

Un index constă dintr-un set de pagini, numite noduri de index, organizate într-o structură arborescentă echilibrată (**B-Tree**), unde numărul de ramuri din dreapta și stânga nodului rădăcină este aproximativ egal. Această structură permite preluarea rapidă și eficientă a informațiilor.



Tipuri de indecși în SQL Server:

1) Index Clusterizat (Clustered Index):

- Stochează rândurile reale de date în nodurile sale frunză (Leaf nodes).
- Datele din tabel sunt fizic aranjate conform ordinii impuse de indexul clusterizat.
- Fiecare tabel poate avea un singur index clusterizat.

2) Index Non-Clusterizat (Nonclustered Index):

- Un index non-clusterizat la nivelul frunzelor conține valorile sortate ale coloanei indexate și un **row locator** – un pointer către restul datelor. Dacă tabelul are un index clusterizat (tabel clusterizat), row locator este **cheia indexului clusterizat**, iar dacă tabelul nu are un index clusterizat (tabel heap), row locator este **row ID** (un pointer către un anumit rând de date dintr-o anumită tabelă).
- Un tabel poate avea mai mulți indecși nonclusterizați.

Index clusterizat și non-clusterizat:

SELECT * FROM Persons
WHERE MobileNo = 111;

NonClustered Key

MobileNo	Pageld
111	1
115	2
119	3

MobileNo	RollNo
111	1
112	2
113	3
114	4

Pageld: 1

MobileNo	RollNo
115	5
116	6
117	7
118	8

Pageld: 2

MobileNo	RollNo
119	9
120	10
121	11
122	12

Pageld: 3

Clustered Key

RollNo	Pageld
1	10
5	11
9	12

RollNo	Name	Age	MobileNo
1	Amit	10	111
2	Abhay	10	112
3	Naveen	12	113
4	Mayank	10	114

Pageld: 10

RollNo	Name	Age	MobileNo
5	Ashish	10	115
6	Shruti	11	116
7	Mateen	12	117
8	Umesh	11	118

Pageld: 11

RollNo	Name	Age	MobileNo
9	Om	13	119
10	Sandeep	10	120
11	Tarun	12	121
12	Deepak	13	122

Pageld: 12

SELECT * FROM Persons
WHERE id = 1;

Crearea (implicită) a Index-ului:

1) PRIMARY KEY

```
CREATE TABLE Persons (  
    Id INTEGER PRIMARY KEY,  
    Name VARCHAR(255)  
);
```

SQL Server creează implicit un **index clusterizat** pentru col. Id

```
CREATE TABLE Persons (  
    Id INTEGER PRIMARY KEY NONCLUSTERED,  
    Name VARCHAR(255)  
);
```

SQL Server creează un **index nonclusterizat** pentru coloana Id (datele nu vor fi ordonate după id)

2) UNIQUE

```
CREATE TABLE Persons (  
    Id INTEGER,  
    Name VARCHAR(255) UNIQUE  
);
```

SQL Server creează implicit un **index nonclusterizat** pentru coloana Name

```
CREATE TABLE Persons (  
    Id INTEGER,  
    Name VARCHAR(255) UNIQUE CLUSTERED  
);
```

SQL Server creează un **index clusterizat** pentru coloana Name (datele vor fi ordonate după name)

CREATE INDEX:

```
CREATE INDEX index_name  
ON table_name (column_name)
```

Se creează implicit un index nonclusterizat neunical.

```
CREATE UNIQUE CLUSTERED INDEX ix_name  
ON table_name (column_name)
```

Se creează un index clusterizat unical.

```
CREATE UNIQUE INDEX index_name  
ON table_name (column_name)
```

Se creează un index nonclusterizat unical.

```
CREATE CLUSTERED INDEX index_name  
ON table_name (column_name)
```

Se creează un index clusterizat neunical.

```
CREATE INDEX index_name  
ON table_name (column_name)  
WHERE [filter-condition]
```

Se creează un index nonclusterizat filtrat.

```
CREATE INDEX idx_SalesTotal  
ON Orders (SalesTotal)  
WHERE SalesTotal > 1000.00;
```

Exemplu de index filtrat.

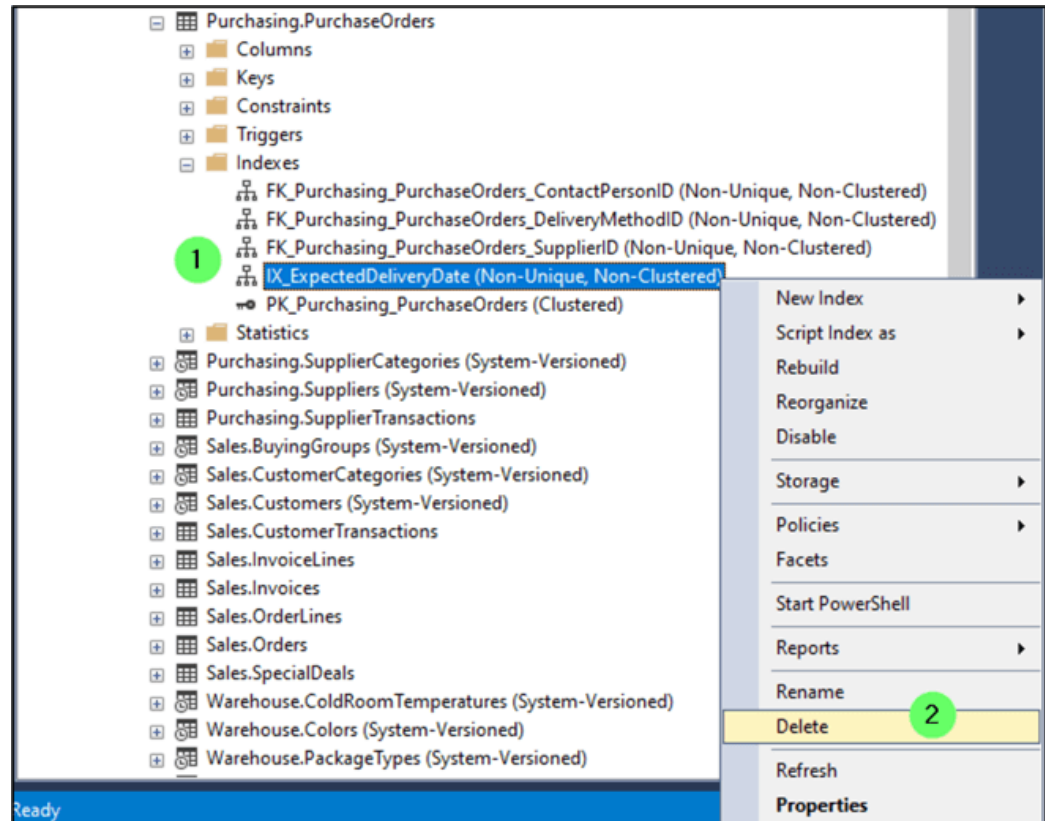
DROP INDEX:

DROP INDEX `index_name`
ON `table_name`

DROP INDEX `idx_UniqueEmployeeName`
ON `Employees`;

Exemplu de ștergere a index-ului.

În SQL Server Management Studio (SSMS):



Sarcină 1:

1. Importează baza de date **Les11_Indexes**.

Hint: Pentru a importa baza de date:

- Fă clic dreapta pe Databases din Object Explorer și selectează Attach....
- În fereastra care se deschide, apasă pe Add... și navighează la locația unde se află fișierele **Les11_Indexes.mdf** și **Les11_Indexes_log.ldf** (C:\Program Files\Microsoft SQL Server\MSSQL16.SQLEXPRESS\MSSQL\DATA). La aceste fișiere în properties trebuie să se atribuie permissions Full Control pentru user MSSQL\$SQLEXPRESS.

2. Pentru tabela **test_table** (cu 30000 de rânduri) din baza de date **Les11_Indexes** execută interogarea pentru a găsi înregistrarea cu **rand_integer** = 793295.

Determină timpul cheltuit pentru această interogare.

Hint: Folosește comanda **SET STATISTICS TIME ON** înainte de interogare pentru a activa afișarea timpului de execuție în Messages.

3. Creează un index pe coloana **rand_integer** pentru a îmbunătăți performanța interogărilor. Execută din nou interogarea de la punctul 2 și observă schimbarea timpului de execuție.

4. Elimină indexul creat anterior și testează din nou timpul de execuție al aceleiași interogări.