

PostgreSQL - INDEXES

Indexes are special lookup tables that the database search engine can use to speed up data retrieval. Simply put, an index is a pointer to data in a table. An index in a database is very similar to an index in the back of a book.

For example, if you want to reference all pages in a book that discusses a certain topic, you have to first refer to the index, which lists all topics alphabetically and then refer to one or more specific page numbers.

An index helps to speed up SELECT queries and WHERE clauses; however, it slows down data input, with UPDATE and INSERT statements. Indexes can be created or dropped with no effect on the data.

Creating an index involves the CREATE INDEX statement, which allows you to name the index, to specify the table and which column or columns to index, and to indicate whether the index is in ascending or descending order.

Indexes can also be unique, similar to the UNIQUE constraint, in that the index prevents duplicate entries in the column or combination of columns on which there's an index.

The CREATE INDEX Command

The basic syntax of **CREATE INDEX** is as follows –

```
CREATE INDEX index_name ON table_name;
```

Index Types

PostgreSQL provides several index types: B-tree, Hash, GiST, SP-GiST and GIN. Each Index type uses a different algorithm that is best suited to different types of queries. By default, the CREATE INDEX command creates B-tree indexes, which fit the most common situations.

Single-Column Indexes

A single-column index is one that is created based on only one table column. The basic syntax is as follows –

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

Multicolumn Indexes

A multicolumn index is defined on more than one column of a table. The basic syntax is as follows –

```
CREATE INDEX index_name  
ON table_name (column1_name, column2_name);
```

Whether to create a single-column index or a multicolumn index, take into consideration the column(s) that you may use very frequently in a query's WHERE clause as filter conditions.

Should there be only one column used, a single-column index should be the choice. Should there be two or more columns that are frequently used in the WHERE clause as filters, the multicolumn index would be the best choice.

Unique Indexes

Unique indexes are used not only for performance, but also for data integrity. A unique index does not allow any duplicate values to be inserted into the table. The basic syntax is as follows –

```
CREATE UNIQUE INDEX index_name  
on table_name (column_name);
```

Partial Indexes

A partial index is an index built over a subset of a table; the subset is defined by a conditional expression (called the predicate of the partial index). The index contains entries only for those table rows that satisfy the predicate. The basic syntax is as follows –

```
CREATE INDEX index_name  
on table_name (conditional_expression);
```

Implicit Indexes

Implicit indexes are indexes that are automatically created by the database server when an object is created. Indexes are automatically created for primary key constraints and unique constraints.

Example

The following is an example where we will create an index on COMPANY table for salary column –

```
# CREATE INDEX salary_index ON COMPANY (salary);
```

Now, let us list down all the indices available on COMPANY table using **\d company** command.

```
# \d company
```

This will produce the following result, where *company_pkey* is an implicit index, which got created when the table was created.

Table "public.company"		
Column	Type	Modifiers
id	integer	not null
name	text	not null
age	integer	not null
address	character(50)	
salary	real	
Indexes:		
"company_pkey" PRIMARY KEY, btree (id)		
"salary_index" btree (salary)		

You can list down the entire indexes database wide using the `\di` command –

The DROP INDEX Command

An index can be dropped using PostgreSQL **DROP** command. Care should be taken when dropping an index because performance may be slowed or improved.

The basic syntax is as follows –

```
DROP INDEX index_name;
```

You can use following statement to delete previously created index –

```
# DROP INDEX salary_index;
```

When Should Indexes be Avoided?

Although indexes are intended to enhance a database's performance, there are times when they should be avoided. The following guidelines indicate when the use of an index should be reconsidered –

- Indexes should not be used on small tables.
- Tables that have frequent, large batch update or insert operations.
- Indexes should not be used on columns that contain a high number of NULL values.
- Columns that are frequently manipulated should not be indexed.