Indexes

NGUYEN Hong Phuong
phuongnh@soict.hut.edu.vn

1

Contents

- Introduction
- Index types
- Multicolumn indexes

2

1. Introduction

- A common way to enhance database performance
- Suppose we have a table:

```
CREATE TABLE test1(
    id integer,
    content varchar
);
```

- Query:

```
SELECT content FROM test1 WHERE id =
    233;
```

3

1. Introduction (cont)

- System will scan entire table, row by row, to find all matching entries.
- Only a few rows will be returned
- This is clearly an inefficient method
- => Maintain an index on id column, it may only have to walk a few levels deep into a search tree.

4

An example of the proceedings

- The proceedings always have author index at the end of them.
- A reader can quickly scan the author index to find out an author whom he/she cares, and he/she can flip to the appropriate pages. He/she does not need to read the entire material.
- Task: anticipate/foresee the items that readers are likely to look up.

5

Create and remove index

- CREATE INDEX test1_id_index ON test1(id);
- DROP INDEX test1_id_index
- Creating an index on a large table can take a long time.

6

2. Index types

- PostGreSQL: B-tree, Hash, GiST and GIN
- Each index type uses a different algorithm that is most suitable to different type of queries.
- Default: CREATE INDEX command (B-tree).

7

2. Index types

- B-trees can handle equality and range queries on data that can be sorted into some ordering. PostgreSQL query planner will consider using a B-tree index whenever an indexed column is involved in a comparison using one of these operators: <,<=,=,>=,>, BETWEEN, IN, IS NULL
- Hash index can only handle simple equality comparisons

8

2. Index types

- When using hash index?
  - An indexed column is involved in a comparison using = operator
  - CREATE INDEX name ON table USING hash (column);
- GiST indexes: many different indexing strategies can be implemented.
  - <<, >>, @>, @<, &&,…

9

2. Index types

- GIN indexes are inverted indexes which can handle values that contain more than one key
  - <@, @>, =, &&

10

3. Multicolumn indexes

- An index can be defined on more than one column of a table

```
CREATE TABLE test2(
    major int,
    minor int,
    name varchar);
```

  - SELECT name FROM test2 WHERE major = const1 AND minor = const2;
  - CREATE INDEX test2_mm_idx ON test2 (major, minor);

11

3. Multicolumn indexes

- Currently, only the B-tree and GiST index types support multicolumn indexes (up to 32 column)

12