

Slide 10 Intro to Indexing + SQL Indexing

CSF2600700 - BASIS DATA SEMESTER GENAP 2019/2020 This slide is a modification to supplementary slide of "Database System", 7th edition, Elmasri/Navathe, 2015: Chapter 17 Indexing Structures for Files and Physical Database Design



- o Indexing is a way of sorting a number of records on multiple fields/attributes.
- Indexes are special lookup tables that the database search engine can use to <u>speed up</u> <u>data retrieval</u>, without reading the whole table.
- Simply, an index is a pointer to data in a table.
- They are based upon one or more columns but stored as a separate entity.



An index in a database is very similar to an index in the back of a book.

Index Dial type 4, 12 About cordless telephones 51 Directory 17 Advanced operation 17 DSL filter 5 Answer an external call during an intercom call 15 Edit an entry in the directory 20 Answering system operation 27 Edit handset name 11 Basic operation 14 FGG, AGTA and IG regulations 53 Battery 9, 38 Find handset 16 Gall log 22, 37 Call waiting 14 Handset display screen messages 36 Handset layout 6 Chart of characters 18 Date and time 8 Important safety instructions 39 Delete from redial 26 Index 56-57 Delete from the call log 24 Installation 1 Delete from the directory 20 Install handset battery 2 Delete your announcement 32 Intercom call 15 Desk/table bracket installation 4 Internet 4 Dial a number from redial 26



An index file consists of records (called index entries) of the form

search-key pointer

- Search Key: attribute to set of attributes used to look up records in a file.
- Pointer: pointer to the record or the block
- Index files are typically much smaller than the original file. Why?
- An index is an auxiliary file that makes it more efficient to search for a record in the data file.



- An index helps speed up <u>SELECT</u> queries and <u>WHERE</u> clauses, but it slows down data input, with <u>UPDATE</u> and <u>INSERT</u> statements. Why?
- Indexes can be created or dropped with <u>no</u> effect on the data.



Types of Index

- Primary Indexes vs. Secondary Indexes
 - Primary index: in a sorted organization file, the index whose search key is the <u>same as the sequential</u> order of the file.
 - Secondary index: an index whose search key is different from the sequential order of the file. Secondary indexes provide a mechanism for specifying an additional key for a base relation that can be used to retrieve data more efficiently



Types of Index

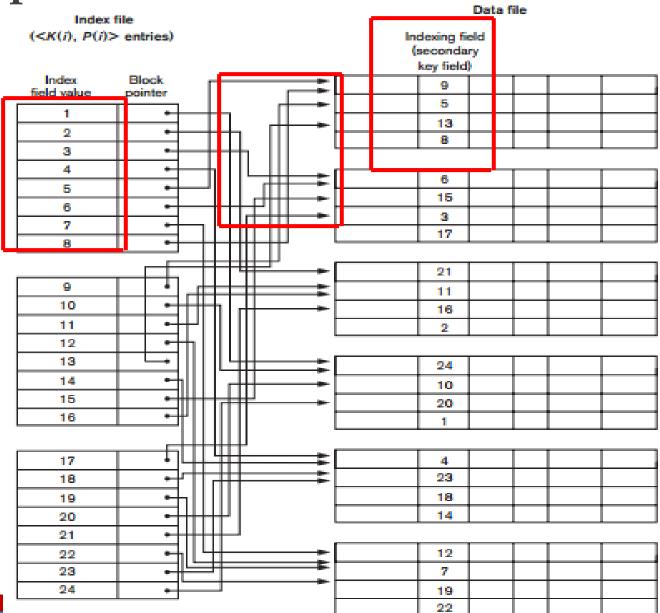
- Sparse vs. Dense Indexes
 - Sparse index: index on attribute for sorting the file, only one index entry per block of data record.
 - Dense index: one index entry per data record.



Example of INDEX Figure 18.1 (Primary key field) Primary index on the ordering key field of the file shown in Figure 17.7. Name Birth date Job Salary Sex Aaron, Ed Abbot, Diane Primary, sparse index ŧ Acosta, Marc Adams, John Adams, Robin ŧ Akers, Jan Index file Alexander, Ed $(\langle K(I), P(I) \rangle$ entries) Alfred, Bob Block anchor Allen, Sam primary key Block pointer Aaron, Ed Allen, Troy Adams, John Anders, Keith Alexander, Ed ŧ Allen, Troy Anderson, Rob Anderson, Zach Anderson, Zach Arnold, Mack Angel, Joe Archer, Sue Arnold, Mack Arnold, Steven Atkins, Timothy Wong, James Wood, Donald Wong, James I Wright, Pam Woods, Manny Wright, Pam Wyatt, Charles Zimmer, Byron

Example of INDEX

Secondary, dense index





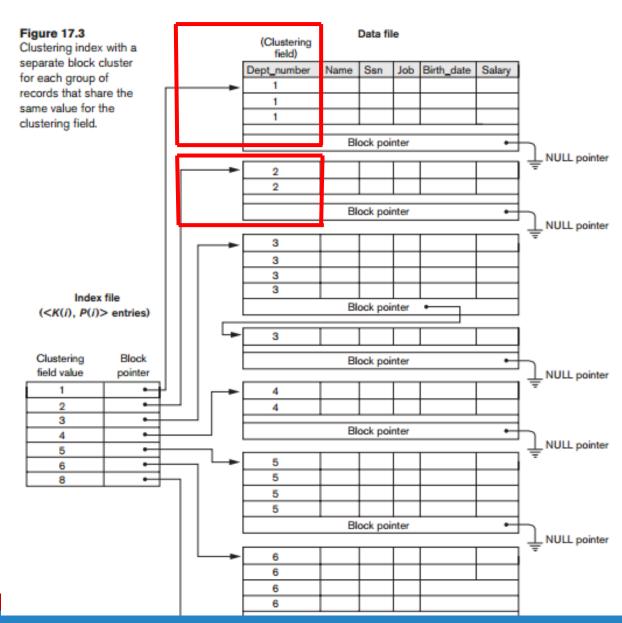
Types of Index

- Clustering Indexes
 - If the ordering field is <u>not UNIQUE</u> or <u>non-key</u>, such as DEPT_NUMBER in EMP table, we can create a clustering index.
 - In the index, one index entry for each distinct value



Example of INDEX

Clustering index



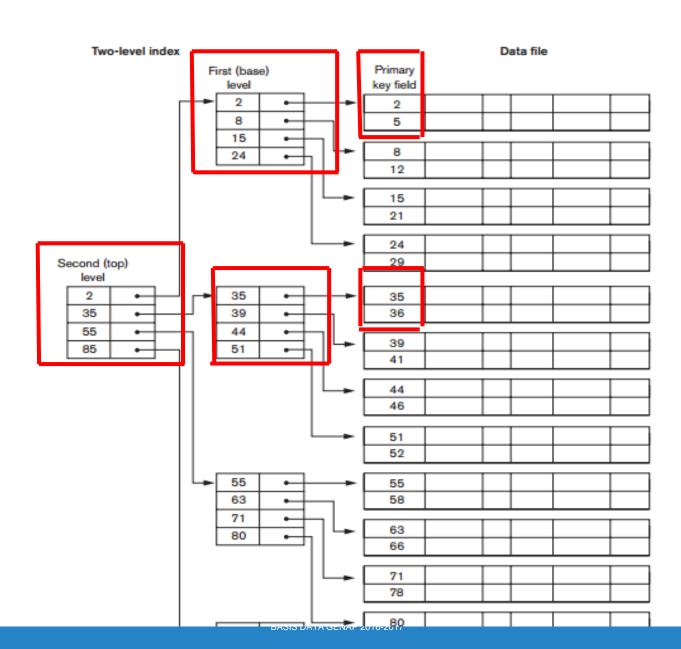
Types of Index

- Single-Level vs. Multi-Level Indexes
 - One index per file
 - Multiple index per file. Index of index.



Example of INDEX

Two-level index





Types of Index

Table 18.1 Types of Indexes Based on the Properties of the Indexing Field

- ·	=	_
	Index Field Used for Physical Ordering of the File	Index Field Not Used for Physical Ordering of the File
Indexing field is key	Primary index	Secondary index (Key)
Indexing field is nonkey	Clustering index	Secondary index (NonKey)



Choose indexes – Guidelines for choosing 'wish-list'

- 1. Do not index small relations.
- 2. Index PK of a relation if it is not a key of the file organization.
- 3. Add secondary index to a FK if it is frequently accessed.
- 4. Add secondary index to any attribute heavily used as a secondary key.
- 5. Add secondary index on attributes involved in: selection or join criteria; ORDER BY; GROUP BY; and other operations involving sorting (such as UNION or DISTINCT).



Choose indexes - Guidelines for choosing 'wish-list'

- 6. Add secondary index on attributes involved in built-in functions.
- 7. Avoid indexing an attribute or relation that is frequently updated.
- 8. Avoid indexing an attribute if the query will retrieve a significant proportion of the relation.
- 9. Avoid indexing attributes that consist of long character strings.





SQL Indexing

INDEX

Creating an index involves the <u>CREATE INDEX</u> statement,

- Which allows you to name the index,
- To specify the table and which column(s) to index,
- And to indicate whether the index is in ascending or descending order

The basic syntax:

- oCREATE INDEX index_name ON
 table_name (column name);
- •Ex: CREATE INDEX Emp_idx ON EMPLOYEE (Ssn)



INDEX

Multiple column index

```
CREATE INDEX index_name ON table_name (column_name1, column_name2);

Ex: CREATE INDEX WORKS ON Idx
```

Ex: CREATE INDEX WORKS_ON_Idx ON WORKS_ON (pno, hours)

Which column to choose is based on frequently queried column in WHERE clause



DROP INDEX

DROP INDEX index name;

Care should be taken when dropping an index because performance may be slowed or improved



WHEN NOT TO USE INDEX

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.



Latihan

- 1. Mengapa kita perlu menghindari pembuatan index pada kolom yang sering dilakukan operasi UPDATE atau INSERT?
- 2. Jika kita menghapus index, maka data pada kolom yang indexnya dihapus juga akan ikut terhapus. Benar atau salah pernyataan ini? Berikan alasannya.



Latihan (2)

3. Apa tipe index dari gambar ini? Jelaskan alasannya

Fact Table

Index Table

Secondary Key (Town)	Customer Reference (ID)
Chicago	ID: 5
Chicago	ID: 9
Chicago	ID: 1000
Portland	ID: 3
Portland	ID: 7
Redmond	ID: 1
Redmond	ID: 4
Redmond	ID: 6
Redmond	ID: 8
Seattle	ID: 2

(Primary Key (Customer ID)	Customer Data
1	1	LastName: Smith, Town: Redmond,
4	2	LastName: Jones, Town: Seattle,
4E	3	LastName: Robinson, Town: Portland,
1	4	LastName: Brown, Town: Redmond,
-	5	LastName: Smith, Town: Chicago,
,	6	LastName: Green, Town: Redmond,
-	7	LastName: Clarke, Town: Portland,
,	8	LastName: Smith, Town: Redmond,
4	9	LastName: Jones, Town: Chicago,
4	1000	LastName: Clarke, Town: Chicago,
Г		

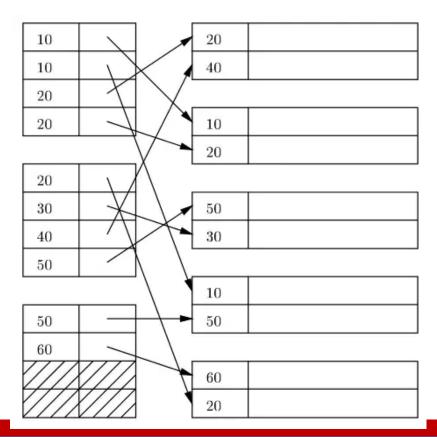
Index Table

	Secondary Key (LastName)	Customer Reference (ID)
1	Brown	ID: 4
1	Clarke	ID: 7
d	Clarke	ID: 1000
ł	Green	ID: 6
N	Jones	ID: 2
4	Jones	ID: 9
l		
N	Robinson	ID: 3
N	Smith	ID: 1
١	Smith	ID: 5
N	Smith	ID: 8



Latihan (3)

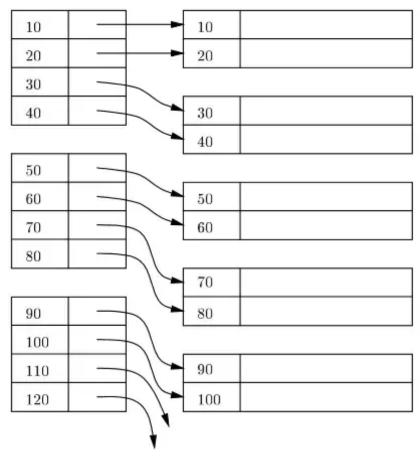
4. Apa tipe index dari gambar ini? Jelaskan alasannya.





Latihan (4)

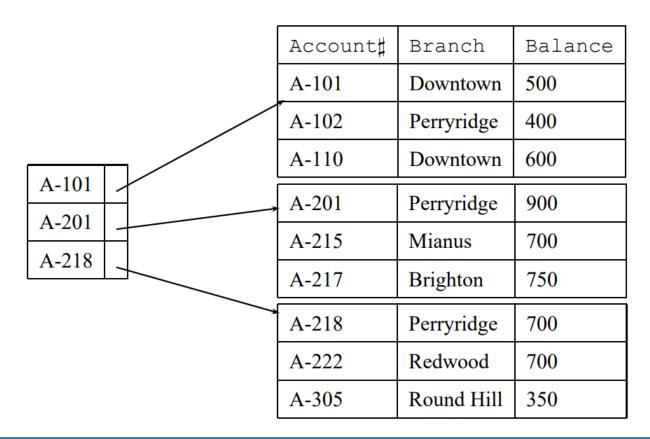
5. Apa tipe index dari gambar ini? Jelaskan alasannya.





Latihan (5)

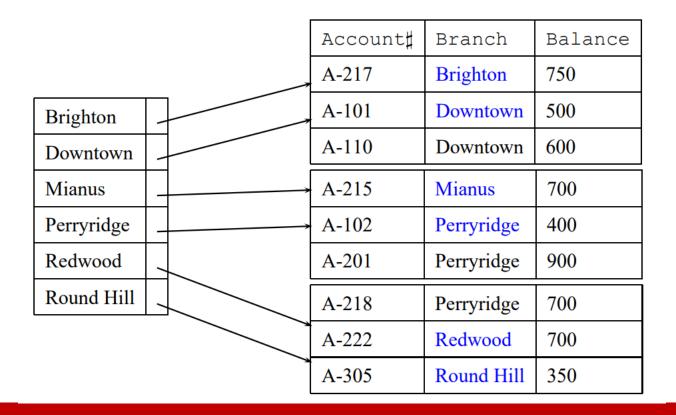
6. Apa tipe index pada gambar ini? Jelaskan alasannya.





Latihan (6)

7. Apa tipe index pada gambar ini? Jelaskan alasannya.





Latihan (7)

8. Sebuah tabel bernama DEPARTMENT tersusun dari tiga kolom yaitu Dno, Dname dan DLocation. Kolom/atribut yang sering digunakan sebagai query statement pada tabel DEPARTMENT tersebut adalah Dno dan DLocation. Buatlah sebuah sql statement untuk membuat index bernama department_idx yang dibuat pada tabel DEPARTMENT tersebut.

