

CMPE321 Introduction to Database Systems
2016/2017-2

Berk Erol

Storage Manager Design

Berk Erol
2014400189

Table of Contents

Introduction	3
Assumptions and Constraints.....	4
System Catalog	4
Type	4
Page	5
Record.....	5
Data Structures	6
System Catalog	6
Type	6
Page	6
Record	7
Operations	8
DDL Operations	8
Create a Type.....	8
Delete a Type.....	8
List all Types	8
DML Operations.....	9
Create a Record	9
Delete a Record	9
Update a Record.....	10
Search for a Record	10
List all Records.....	11

Introduction

This project is about designing a storage manager. This document explains the assumptions and constraints required to manage the system & the data structures (with diagrams) and operations (with pseudocodes) required to store the data in the system. This DBMS enables users to run various DDL and DML operations as the followings

- DDL Operations
 1. Create a type
 2. Delete a type
 3. List all types
- DML Operations
 1. Create a record
 2. Delete a record
 3. Update a record
 4. Search for a record (by key field)
 - < views all records whose key fields are less than a specific value.
 - > views all records whose key fields are more than a specific value.
 - = views all records whose key fields are equal to a specific value.
 5. List all records of a type

Assumptions and Constraints

System Catalog

- There is only one system catalog.
- Creating or deleting system catalogs are not allowed.
- System Catalog file shall have the following format: SystemCatalog.txt
- System Catalog header
 - Name of the storage system (50 bytes)
 - # of types (5 bytes)
- Storage system name
 - Storage system name shall be alphanumeric.
 - The length of storage system name shall be between 3 and 50 bytes.
- System Catalog shall have at least one type and at most 99999 types.

Type

- Each type is stored in a different file.
- Each file shall be able to store all pages for a type.
- File shall have the following format: <TypeName>.txt
- Type name
 - Each type name shall be unique.
 - Type names shall be alphanumeric.
 - The length of type names shall be between 3 and 30 bytes.
- File header
 - Usage status (1 byte)
 - # of pages (2 bytes)
 - # of fields (2 bytes)
 - Field names (160 bytes)
- Field name
 - Each field name shall be unique.
 - Field names shall be alphanumeric.
 - The length of field names shall be between 3 and 10 bytes.
- Size of each type shall be at most 101541 bytes.
- Each type shall have at most 99 pages. $(101541 - 1 - 2 - 2 - 160) / 1024$
- Each type shall have at most 1485 records. $(99 * 15)$
- Each type shall have at most 23760 fields. $(1485 * 16)$
- Types cannot be deleted physically (only their status change).

Page

- Each page stores records.
- Page header
 - Page number (2 bytes)
 - # of active records (2 bytes)
 - # of deleted records (2 bytes)
- Line endings $(15+1) * 2 = 32$ bytes.
- Unused space 11 bytes.
- Size of each page shall be at most 1024 bytes.
- Each page shall have at most 15 records. $(1024-2-2-2-32-11)/65$
- Each page shall have at most 240 fields. $(15*16)$
- Pages do not need to be deleted physically because records are not deleted physically.

Record

- Each record stores fields.
- Record header
 - Usage status (1 byte)
- Key field
 - Each key field shall be unique.
 - There shall be one and only one key field in a record.
 - Key field shall always be the first field in its record.
- Size of each record shall be at most 65 bytes.
- Each record shall have at most 16 fields. $(65-1)/4$
- Records cannot be deleted physically (only their status change).
- Size of each field shall be at most 4 bytes.
- Each field stores data as 32 bit integers.
- Fields can be deleted physically (except key field).

Data Structures

System Catalog

It is the main file of the storage system. It keeps all general information about the storage system and its files. Users first reach this file if they want to make an operation in the storage system. There is only one system catalog. Creating or deleting system catalogs are not allowed.

System Catalog

System Catalog Header
Type Name 1
Type Name 2
Type Name ...
Type Name n

System Catalog Header

System name	# of types
-------------	------------

Type

The storage system consists of several types. Each type is stored in a different file. Each file stores all pages for a type. Usage status specifies whether the type is deleted or not (1 for active, 0 for deleted). Types cannot be deleted physically (only their status change). This property speeds up the type deletion process. Types can have different number of pages, records, fields and different sizes but the maximum amount for these values are the same for all types as specified in Assumptions and Constraints section.

File

File Header
Page 1
Page 2
Page ...
Page 99

File Header

Usage status	# of pages	# of fields	Field Name 1	Field Name ...	Field Name 16
--------------	------------	-------------	--------------	----------------	---------------

Page

Each page stores records. Each page shall have a fixed size of 1024 bytes but a type (file) can have multiple pages. There shall always be enough number of pages in a file for all records of a type (Pages are created after the last page of the type when there is no empty record slot in the last page of the type but pages do not need to be deleted physically because records are not deleted physically). This property speeds up the record creation and deletion processes. Pages can have different number of records and fields and different sizes but the maximum amount for these values are the same for all pages as specified in Assumptions and Constraints section.

Page

Page Header
Record 1 (Status 1)
Record ... (Status 1)
Record 10 (Status 1)
Record 1 (Status 0)
Record ... (Status 0)
Record 5 (Status 0)

Page Header

Page number	# of active records	# of deleted records
-------------	---------------------	----------------------

Record

Each record stores fields. Usage status specifies whether the record is deleted or not (1 for active, 0 for deleted). New records are created at the right position in the used records (to keep used records sorted according to key field). Records cannot be deleted physically (only their status change) and newly deleted records are moved to the end of the file. This way used records remain sorted and deleted records come after used records. These properties speed up the record searching and listing processes. Records can have different number of fields and different sizes but the maximum amount for these values are the same for all records as specified in Assumptions and Constraints section. There is only one key field which is the first field. Each field stores data as integers. Fields can be deleted physically (except key field).

Record
Record Header
Field 1 (Key Field)
Field 2
Field ...
Field 16

Record Header
Usage status