WSB University

Databases and Data Warehouses Project

On

Methods of data recovery and updating in a data warehouse and CRUD operations

Academic year 2020/2021

Student Name: IDA OGAH

Student ID: 45086

Program: MSc Data Science (IT)

Email: ida.ogah@student.wsb.edu.pl

akademia**wsb**.edu.pl

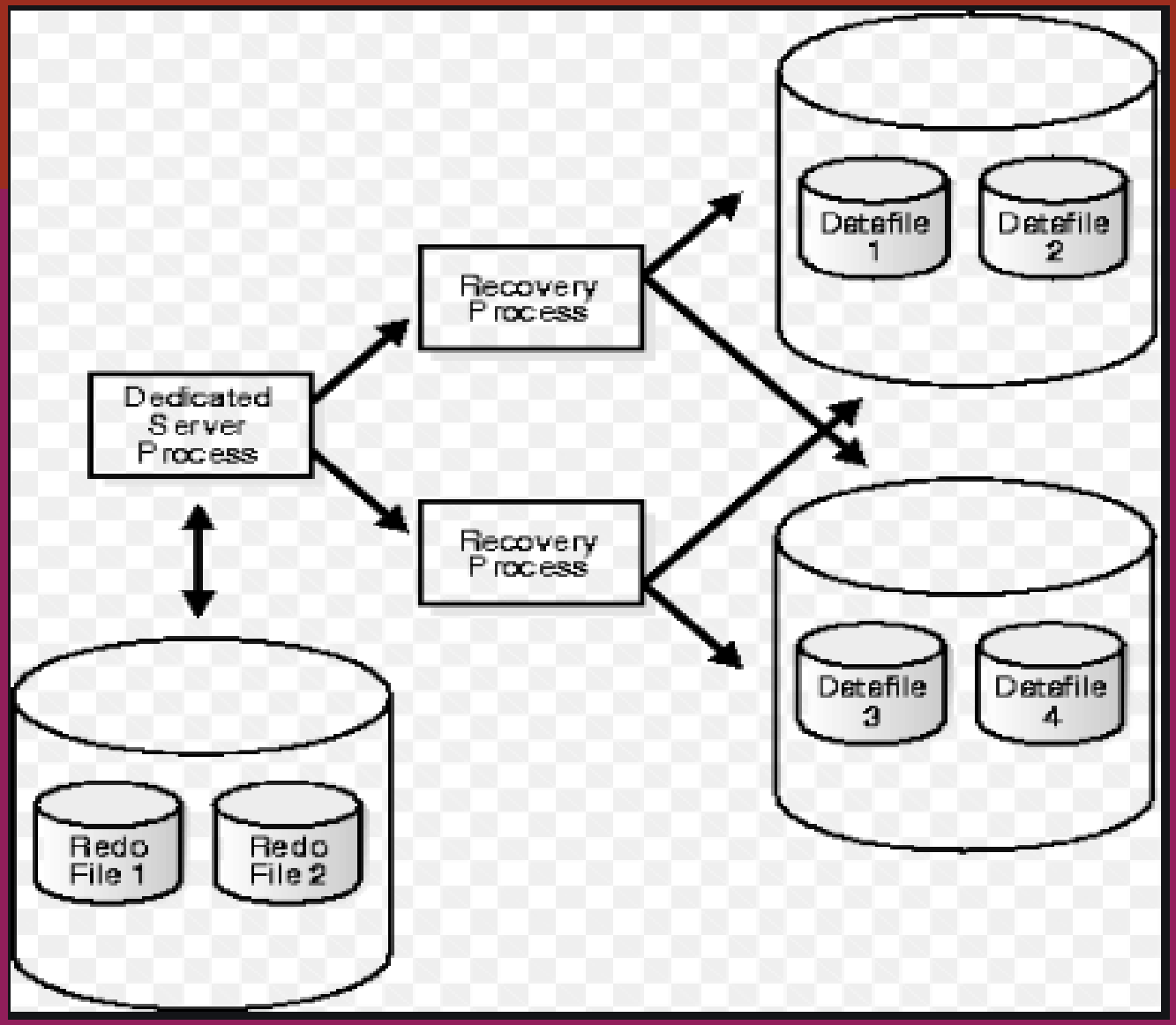
METHODS OF DATA RECOVERY

- **Introduction:**
  - Database Recovery Methods are techniques used to recover the lost data due to system crash, transaction errors, viruses, catastrophic failure, incorrect commands execution, etc.
  - In database systems, like any other computer system, are subject to failures but the data stored in it must be available as and when required.
  - When a database fails it must possess the facilities for fast recovery. It must also have atomicity.
  - There are both automatic and non-automatic ways for both, backing up of data and recovery from any failure situations.

DATA RECOVERY & SYSTEM LOG FILE

- The log file is the system file that keep track of all transactions and operations that affect the values of database items.
- A transaction, T is said to have reach its **commit** point when all its operations that access the database have been executed successfully i.e. the transaction has reached the point at which it will not abort or terminate without completing.

DATABASE RECOVERY



Source: docs.oracle.com

INFORMATION NEEDED FOR TRANSACTION RECOVERY

a. **start_transaction(T):** This log entry records that transaction T starts the execution.
b. **read_item(T, X):** This log entry records that transaction T reads the value of database item X.
c. **write_item(T, X, old_value, new_value):** This log entry records that transaction T changes the value of the database item X from old_value to new_value. The old value is known as a before-image of X, & the new value is known as an afterimage of X.
d. **commit(T):** This log entry records that transaction T has completed all accesses to the database successfully and its effect can be committed (recorded permanently) to the database.
e. **abort(T):** This records that transaction T has been aborted without completing.
f. **checkpoint:** Checkpoint is a mechanism where all the previous logs are removed from the system and stored permanently in a storage disk.

AUTOMATIC RECOVERY / UPDATING

1. **Undoing** – If a transaction crashes, then the recovery manager may undo transactions i.e. reverse the operations of a transaction by examining a transaction log for write_item(T, x, old_value, new_value) and setting the value of item x in the database to old-value.

   Two Types of Undoing Techniques –

   a. **Deferred update** – This technique does not physically update the database on disk until a transaction has reached its commit point. This is also known as the **No-undo/redo algorithm.**
   b. **Immediate update** – The database may be updated by some operations of a transaction before the transaction reaches its commit point. This is also known as **undo/redo algorithm.**

AUTOMATIC RECOVERY / UPDATING CONT.

2. **Caching/Buffering** – In this one or more disk pages that include data items to be updated are cached into main memory buffers and then updated in memory before being written back to disk.

3. **Shadow paging** – It provides atomicity and durability. A directory with n entries is constructed, where the ith entry points to the ith database page on the link.

DATA & DATABASE BACKUP TECHNIQUES

1. Full database backup – In this full database including data and database, Meta information needed to restore the whole database, including full-text catalogs are backed up in a predefined time series.

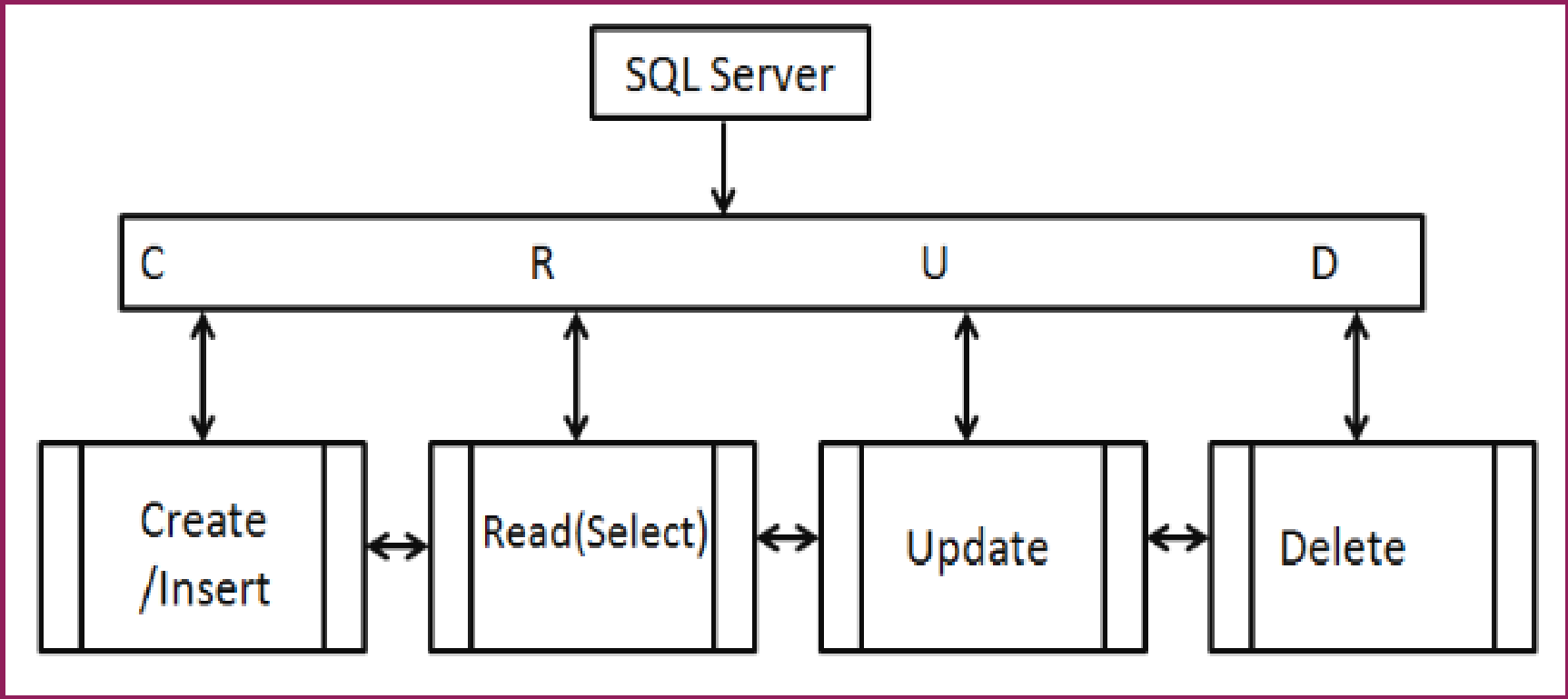
2. Differential backup – It stores only the data changes that have occurred since last full database backup. When same data has changed many times since last full database backup, a differential backup stores the most recent version of changed data.

3. Transaction log backup – It is the backup of transaction log entries and contains all transaction that had happened to the database.

CRUD OPERATIONS

- **What is CRUD Operations?**
  - **CRUD operations are foundation operations every database developer and administrator needs to understand.**
  - **In computer programming, CRUD stands for Create, Read, Update, and Delete (CRUD).**

CRUD OPERATIONS IN SQL SERVERS



HOW CRUD OPERATIONS WORK

❖ CREATE

The first letter of CRUD, 'C', refers to CREATE.
this operation is used to add or insert a new record into the database. SQL uses INSERT INTO statement to create new records within the table.

Let us create a simple table named *Demo* for this example.

```
USE AdventureWorks2016;
GO
DROP TABLE IF EXISTS Demo;
CREATE TABLE dbo.Demo
(id   INT,
name VARCHAR(100)
);
```

To insert multiple rows, follow the below syntax

```
INSERT INTO <tablename> (column1,column2,….)
VALUES(value1,value2,…. ),( value1,value2,…. ), (value1,value2,…. )…
```

HOW CRUD OPERATIONS WORK CONT.

- **READ**

The second letter of CRUD, 'R', refers to READ (SELECT). The word 'read' retrieves data or record-set from a listed table(s). SQL uses the SELECT command to retrieve the data.

This commands can be used in SQL Server Management Studio or SQL Server Data Tools or sqlcmd, based on preference.

For example, to read related data from the specified table, refer to the below syntax.

```
SELECT * FROM <TableName>
```

HOW CRUD OPERATIONS WORK CONT.

- **UPDATE**

The third letter of CRUD, 'U', refers to UPDATE.
In SQL, using the UPDATE keyword brings a change to the selected record(s) of the table.

The basic syntax for an update:

```
UPDATE <TableName>
SET Column1=Value1, Column2=Value2,…
WHERE <Expression>
```

HOW CRUD OPERATIONS WORK CONT.

- **DELETE**

The fourth letter of CRUD, 'D', refers to DELETE.
In SQL, DELETE command is use to remove or delete the selected record(s) from the table.

For example, to delete related data from the specified table, refer to the below syntax

```
DELETE FROM <TableName>
WHERE <Expression>
```

SUMMARY & CONCLUSION

In conclusion, implementation of the CRUD operations (Create, Read, Update, Delete, and Insert) in most cases are done by database managers to modify data and databases.

On the other hand, data recovering and update operations can be implemented automatically through system log file or database managers.

REFERENCES:

1. Asuncion, A. & Newman, D.J. (2010). UCI machine learning repository. *http://archive.ics.uci.edu/ml*.
2. Chung, S.M. & Mangamuri, M. (2005). Mining association rules from the star schema on a parallel NCR teradata database system.
3. Himanshi, (2019). **Database Recovery Techniques in DBMS.** *https://www.geeksforgeeks.org/database-recovery-techniques-in-dbms*.
4. Prashanth Jayaram, (2018). CRUD operations in SQL Server. *https://www.sqlshack.com/crud-operations-in-sql-server/*.

Akademia WSB

THANK YOU FOR YOUR ATTENTION