DBMS Day 2

RCD - Rapidly Changing Dimension

SCD - Slowly Changing Dimension

UCD – Unchanging Changing Dimension

**Demonstrating SCD Type1,Type2,Type3 using SQL**

CREATE TABLE EMPLOYEE(

EmployeeName varchar(255),

EmployeeDOB date,

EmployeeGender varchar(15),

EmployeeLocation varchar(15),

EmployeeContact number,

PRIMARY KEY (EmployeeContact)

);

CREATE TABLE D\_EMPLOYEE(

PK\_Employee\_Key int,

E\_Contact\_number number,

E\_Name varchar(255),

E\_DOB date,

E\_Gender varchar(15),

E\_Location varchar(15),

Created\_date timestamp,

Updated\_date timestamp,

PRIMARY KEY (PK\_Employee\_Key)

);

INSERT INTO EMPLOYEE (EmployeeName, EmployeeDOB, EmployeeGender, EmployeeLocation, EmployeeContact)

VALUES ('Ram Kumar', '01-JAN-1990', 'Male', 'Delhi', 7898528527);

INSERT INTO EMPLOYEE (EmployeeName, EmployeeDOB, EmployeeGender, EmployeeLocation, EmployeeContact)

VALUES ('Sunita', '02-FEB-1991', 'Female', 'Bangalore', 9517538525);

INSERT INTO EMPLOYEE (EmployeeName, EmployeeDOB, EmployeeGender, EmployeeLocation, EmployeeContact)

VALUES ('Ajay', '01-JAN-1985', 'Male', 'Bangalore', 9638527415);

INSERT INTO EMPLOYEE (EmployeeName, EmployeeDOB, EmployeeGender, EmployeeLocation, EmployeeContact)

VALUES ('Nitin', '01-FEB-1985', 'Male', 'Chennai', 8527539515);

INSERT INTO D\_EMPLOYEE (PK\_Employee\_Key, E\_Contact\_number, E\_Name, E\_DOB, E\_Gender, E\_Location, Created\_date, Updated\_date)

VALUES

(1, 7898528527, 'Ram Kumar', '01-JAN-1990', 'Male', 'Delhi', '30-Aug-2023 10:02', '31-Aug-2023 11:00');

INSERT INTO D\_EMPLOYEE (PK\_Employee\_Key, E\_Contact\_number, E\_Name, E\_DOB, E\_Gender, E\_Location, Created\_date, Updated\_date)

VALUES

(2, 9517538525, 'Sunita', '02-FEB-1991', 'Female', 'Bangalore', '30-Aug-2023 10:02', '');

INSERT INTO D\_EMPLOYEE (PK\_Employee\_Key, E\_Contact\_number, E\_Name, E\_DOB, E\_Gender, E\_Location, Created\_date, Updated\_date)

VALUES

(3, 9638527415, 'Ajay', '01-JAN-1985', 'Male', 'Bangalore', '30-Aug-2023 10:02', '');

INSERT INTO D\_EMPLOYEE (PK\_Employee\_Key, E\_Contact\_number, E\_Name, E\_DOB, E\_Gender, E\_Location, Created\_date, Updated\_date)

VALUES

(4, 8527539515, 'Nitin', '01-FEB-1985', 'Male', 'Chennai', '30-Aug-2023 10:02', '');

select \* from D\_EMPLOYEE;

--- TYPE 1 SLOWLY CHANGING DIMENSION (SCD) :- Updating Date ---

update emplyoyee set EmployeeLocation='Pune' where EmployeeContact=8527539515;

select \* from emplyoyee;

update d\_emplyoyee set E\_Location='Pune', Updated\_date = '31-Aug-2023 11:30' where E\_Contact\_number = 8527539515;

select \* from D\_EMPLYOYEE;

--- TYPE 2 SLOWLY CHANGING DIMENSION (SCD) :- Keeping History of Change (Version) ---

CREATE TABLE D\_EMPLOYEE\_scd2(

PK\_Employee\_Key int,

E\_Contact\_number number,

E\_Name varchar(255),

E\_DOB date,

E\_Gender varchar(15),

E\_Location varchar(15),

start\_date timestamp,

end\_date timestamp,

version int,

flag varchar(2),

PRIMARY KEY (PK\_Employee\_Key)

);

INSERT INTO D\_EMPLOYEE\_scd2 VALUES (1, 7898528527, 'Ram Kumar', '01-JAN-1990', 'Male', 'Delhi', '29-Aug-2023 01:00', '',1,'Y');

INSERT INTO D\_EMPLOYEE\_scd2 VALUES (2, 9517538525, 'Sunita', '02-FEB-1991', 'Female', 'Bangalore', '29-Aug-2023 01:00', '',1,'Y');

INSERT INTO D\_EMPLOYEE\_scd2 VALUES (3, 9638527415, 'Ajay', '01-JAN-1985', 'Male', 'Bangalore', '29-Aug-2023 01:00', '',1,'Y');

INSERT INTO D\_EMPLOYEE\_scd2 VALUES (4, 8527539515, 'Nitin', '01-FEB-1985', 'Male', 'Chennai', '29-Aug-2023 01:00', '',1,'Y');

Update D\_EMPLOYEE\_scd2 set end\_date='30-Aug-2023 1:00',flag='N' where PK\_Employee\_Key=1;

INSERT INTO D\_EMPLOYEE\_scd2 VALUES (5, 7898528527, 'Ram Kumar', '01-JAN-1990', 'Male', 'Delhi', '30-Aug-2023 01:00', '',2,'Y');

Update D\_EMPLOYEE\_scd2 set end\_date='31-Aug-2023 1:00',flag='N' where PK\_Employee\_Key=2;

INSERT INTO D\_EMPLOYEE\_scd2 VALUES (6, 9517538525, 'Sunita', '02-FEB-1991', 'Female', 'Bangalore', '31-Aug-2023 01:00', '',2,'Y');

select \* from D\_EMPLOYEE\_scd2;

--- TYPE 3 SCD 3 : Keeping record of previous data only 1 level back ----

CREATE TABLE D\_EMPLOYEE\_scd3(

PK\_Employee\_Key int,

E\_Contact\_number number,

E\_Name varchar(255),

E\_DOB date,

E\_Gender varchar(15),

E\_Location varchar(15),

Prev\_E\_Contact\_number number,

Prev\_E\_Name varchar(255),

Prev\_E\_DOB date,

Prev\_E\_Gender varchar(15),

Prev\_E\_Location varchar(15),

Created\_date timestamp,

Updated\_date timestamp,

PRIMARY KEY (PK\_Employee\_Key)

);

Insert into D\_EMPLOYEE\_scd3 values (1, 7898528527, 'Ram Kumar', '01-JAN-1990', 'Male', 'Delhi',7898528527, 'Ram Kumar','01-JAN-1990','Male','Bangalore','30-Aug-2023 10:02','31-Aug-2023 11:00');

Select \* from D\_EMPLOYEE\_scd3;

**Big Data**

**A diagram of a process

Description automatically generated**

**HDFS Architecture**

**A diagram of a network

Description automatically generated**

HDFS has a master/slave architecture. An HDFS cluster consists of a single NameNode, a master server that manages the file system namespace and regulates access to files by clients. In addition, there are a number of DataNodes, usually one per node in the cluster, which manage storage attached to the nodes that they run on. HDFS exposes a file system namespace and allows user data to be stored in files. Internally, a file is split into one or more blocks and these blocks are stored in a set of DataNodes. The NameNode executes file system namespace operations like opening, closing, and renaming files and directories. It also determines the mapping of blocks to DataNodes. The DataNodes are responsible for serving read and write requests from the file system's clients. The DataNodes also perform block creation, deletion, and replication upon Instruction from the NameNode

File system - A user or application can create directories and store file inside these

Data Application – HDFS is designed to reliably store very large files across machines in a large cluster. It stores each file as sequence of blocks. The blocks of a file are replicated for fault tolerance.