Data fragmentation

Horizontal, vertical, mixed

- Decomposing a database into multiple smaller units called FRAGMENTS, which are logically related and correct parts

Characteristics of Fragmentation

- Must be complete ,
- must be possible to reconstruct the original database from the fragments.

A relation can be fragmented in three ways:

- Horizontal Fragmentation
- Vertical Fragmentation
- Mixed Fragmentation.

Types of fragmentation

1] Horizontal fragmentation:

- It is a horizontal subset of a relation which contain those of tuples which satisfy selection conditions.
- Specified in the SELECT operation of the relational algebra on single or multiple attributes
- Consider the Employee relation with selection condition (DNO = 5). All tuples satisfy this condition will create a subset which will be a horizontal fragment of Employee relation. σ(Dno=5) Employee.

Types of Horizontal fragmentation

Primary Horizontal Fragmentation:

- Fragmentation of primary relation
- Employee Table is fragmented for Department No.

Derived horizontal fragmentation:

- Fragmentations of the secondary relations that are dependent on the primary relation ; related with Foreign keys.
- ‘WorksInProject’ : Employee and Project relation, can have Employee_Id as its foreign key , and can be fragmented horizontally over EmployeeId for the Various Projects in a department.
- Find all the ProjectIds from WorksInProject where
- EmployeeId=123

Complete horizontal fragmentation

- Horizontal fragments have each and every tuple of the original relation.
- Each tuple of the original relation will belong to at least one partition.
  - Original Relation :Employee
  - Fragment1: σ(Employee Age<=21) Employee
  - Fragment2: σ(18<Employee Age < 65) Employee
  - Fragment3: σ(Employee Age >= 65) Employee
- 100 Rows in Original Table : employee , then the total number of tuples in the above 3 fragments will be either 100 or more than 100.

Disjoint horizontal fragmentation

- No 2 Fragments EVER have common tuples
- Every one tuple of the original relation belongs to 1 fragment
- Original :Employee
  - Fragment1: σ(Employee Age<=18) Employee

Reconstruction of original relation

- To reconstruct the relation *R* from a *complete* horizontal fragmentation, we need to apply the UNION operation to the fragments.
- Original Employee is constructed via:

Employee ← (Fragment1) U (Fragment2) U (Fragment3)

2] Vertical fragmentation:

- Each site may not need all the attributes of a relation. Thus we use **Vertical fragmentation** which divides a relation “vertically” by columns.
- It is a subset of a relation which is created by a subset of columns.
- Consider the Employee relation:
  - A vertical fragment of can be created by keeping the values of Name, Birthdate, Gender , and Salary.
- Because there is no condition for creating a vertical fragment, each fragment must include the primary key attribute of the parent relation Employee. In this way all vertical fragments of a relation are connected.
- PROJECT operation of the relational algebra is used

$$\Pi_{\text{Name, Address, Gender, Salary}}(\text{Employee})$$

Complete vertical fragmentation

- A set of vertical fragments whose projection lists L1, L2, …, Ln include all the attributes in R but share only the primary key of R
- In this case the projection lists satisfy the following two conditions:
- L1 U L2 U ... U Ln = ATTRS (R)
- Li ∩ Lj = PK(R) for any i j, where ATTRS (R) is the set of attributes of R and PK(R) is the primary key of R.

Reconstruction of original relation

- To reconstruct R from complete vertical fragments a OUTER UNION is applied.

3] Mixed fragmentation:

- A combination of Vertical fragmentation and Horizontal fragmentation.
- This is achieved by SELECT-PROJECT operations which is represented by

$$\Pi_{Li}(\sigma_{Ci}(R))$$

- Select name and salary of all Male Employees from Employees relation whose salary =$50,000

Fragmentation aims to improve:

- Reliability
- Performance
- Balanced storage capacity and costs
- Communication costs
- Security