Chapter 2

Desigining& Implementing Data Warehouse

❖ What Is Data Warehouse Design?

- Data warehouse for business intelligence allows your organization to query data obtained from trusted & use the answer to gain a competitive in your business.
- Data warehouse design is the process of building a solution to integrate from data multiple sources & that support analytical reporting & data analysis.
- Data model methodology & system architecture provides a design, basic form, data warehouse that emphasizes core data quality, quantity & ability to support enterprise while data provision needs.
- Data warehouse is a relational database that is designed for query & analysis greater than transaction processing.
- Data warehouse design for a custom software development consulting, satisfying, training companies.
- There are two types of designing in data warehouse.
  1) Logical Design
  2) Physical Design

1) Logical Design For Data Warehouse

- The process of logical designing invoice arranging data into series of logical relationship called entity & attributes.
- Logical design for logical relationship between objects.
- Entity Relationship (ER) modeling technique can be used for logical design of data warehouse.

- Logical design for a data warehouse part of implementation a data warehouse with Microsoft SQL Server.
- Logical design often starts with a conceptual schema & then generates relation structure.
- Entity relationship involves identify the entity, attribute & relationship them
- An entity is a chunk information which map to a table in database.
- An attribute is a part of an entity that map to a column in a database.
- A unique identifier can be used to make sure the data is consistent.
- A logical design is conceptual & abstract you don't deal with the physical implementation yet.
- You deal only with defining the types of information that you need.
- One technique you can use to model your organization logical information requirement & entity relationship modeling.
- During process business requirement analysis, business analyst talks to the user & examine the user model that exists in the user with the co-operation of the user this is formalized into an entity relationship model which essentially forms the logical model.
- It is important to realize that this logical model is based entirely upon the user requirement.
- Once the logical model is complete it is over to DBD (Database Design).
- At this point a decision is made about the database model & the database engine that will be used.

2) Physical Design For Data Warehouse

- During physical design process you convert the data gathered during the logical design phase into a description of the physical database structure.
- Physical design decision is mainly driven by query performance & database maintenance aspects.
- During the logical design phase you define a model for your data warehouse consists of entities, attributes & relationship.
- The entities are linked together using relationship & attributes are used to describe the entity.
- Physical design structure:
  1) **Table Spaces.**
  2) **Table & Partition Table.**
  3) **Views.**
  4) **Integrity Constraints.**
  5) **Dimension.**
- The physical design deal with the affective way of storing & retrieving the data.
- Physical design involves creation of the database object like tables, indexes, column, primary key, foreign key, view & other database object.
- During the physical design process the expected schema into database structure at this time you have to map like:
  1) **Entity to table.**
  2) **Relationship to foreign key constraint.**
  3) **Attributes to column.**
  4) **Primary unique identifier to primary key constraints.**
  5) **Unique identifier to unique constraints.**
- Physical design of warehouse by using the components diagram & deployment diagrams of UML (Unified Modeling Language).

Design By: Mr.Ronak J. Goda

- Physical decision that may overall development time of a data warehouse such as tables, vertical & horizontal partition of a table, ETL process (Extract Transfer Load) & so on.
- Physical design is a important & highly influences the overall performance of the data warehouse & maintenance.
- The physical design procedure of an information warehouse has to be significant & comparative to its physical storage & public presentation.

❖ Design Dimension Table/Fact Table For Data Warehouse

➢ What is Dimension Table?

- A dimension table is a table in a star schema of data warehouse.
- Data warehouse built using dimension table or data model which consists of fact & dimension tables, dimension tables are used to describe they contain dimension key, values & attributes.
- The time dimension would contain every hours, day, week, months, quarter, year that has occurred since you started your business operations.
- Dimensional table typical small from a few to more thousand rows (records), the occasionally dimension can grow.
- Dimension table can be utilize access for the reports & it's about reusability,
- **Example:** Ane- Commerce Company can create a dimensional table with various columns depending on different subject that would like gain

information from like name of person, address, date of order, shipping information, etc.

- A dimensional table has a primary key column that uniquely identifies each dimension record or row.
- The dimension table is associated with fact table using this key.
- The data warehouse organizes descriptive attributes in columns dimensional record.
- Dimension table are referred by fact table using key when extracting a dimension table in a data warehouse a system generated key is used to uniquely identify a row in the dimension.
- Typically any dimension table has a primary that link all the dimension records to the individual primary records.

➢ **What Is Fact Table?**

- In data warehouse a fact table consists of the measurement matrixes as fact table of a business process.
- A fact table in a primary table in a dimensional model.
- A fact table typically has two types of columns those that contain a fact & that are a foreign key to dimension table.
- Fact table provides the usual values that act as independent variable (by which dimensional attributes are analyzed).
- There are following types of fact tables:
  1) Transactional.
  2) Periodic Snapshot.
  3) Temporary Snapshot.
  4) Accumulate Snapshot.
- A fact table is the central table in a start schema of a data warehouse.

- A fact table holds the data to be analyzed & a dimensional data stores the data about a ways in which the data in the fact table can be analyzed.
- The foreign key column allow join with dimensional table & the major columns contain the data that has been analyzed.
- The fact table consists of fact of a particular business process like sales revenue by month of product.
- Facts are also known as measurement or matrix.
- Measurement types:
  1) **Additive.**
  2) **Non- Additive.**
  3) **Semi- Additive.**

➢ **Steps For Designing A Fact Table**

- **Step 1:** Choosing the business process to model.
- **Step 2:** Declare the gain.
- **Step 3:** Choose the dimension.
- **Step 4:** Identify the fact.

❖ **Difference Between Dimension Table & Fact Table**

Parameter	Dimension Table	Fact Table
Definition	Table to the fact table contains descriptive attributes to be used as query constraint.	Measurement or matrix fact about a business process.
Design	Should be word,	Defined by their grain

	descriptive, complete & quality assured.	or its more atomic level.
Key	Foreign key to the fact table.	Primary key in fact is mapped as foreign key to dimension
Storage	Load detail atomic data into dimensional structure.	Help to storage report levels & filter domain values in dimension table.
Task	Collection of references information about a business.	Fact table is a measurable event for which dimension table data is collected & is used for analysis & reporting.
Characteristics	Dimension table connected to the fact table & located at the start schema.	Fact table located at the center of the start schema & surrounded by dimensions.
Types Of Data	Dimension table contain attributes which describe the details of dimension **Ex:** Product dimension can contain product Id, Name, Prize, Description, Category.	Fact table would contain information like sells, again set of dimension like product & date.
Hierarchy	Contain hierarchy **For Ex:** Location would contain info like Address, City, State, Pin code, Country.	Does not contain hierarchy.

❖Design & Implementation Effective Physical Data Structure Or Data Warehouse

- The physical data modeling apply physical constraint such as space performance of data.
- The physical data model is related to database system & data warehouse tools that you will use.
- The purpose of physical data structure to design actual physical implementation.
- Once a data warehouse is implemented & your customer begins using it, they will generate new request & requirements.

❖ Physical Data Design

- Physical design translates the logical data model into a set of SQL statement that defines the database.
- Physical data model database independence & other techniques may be applied relational database engine,
- Physical data models describe how the system will be implemented using a specific DBMS system.
- This model is typically created by DBA developers the purpose of actual designing & implementation of the database.
- Physical data & design structure also help to visualize like columns, key, constraints, indexes, function, procedure & other RDBMS objects.

- The physical data design contains relationship between tables that which address connectivity on established other objects.
- The physical set of table sitting between the operational system & data warehouse or a specially administrative. The data warehouse it.
- The data design should take advantage of the database administrator capability like database partition, multi dimension clustering, table partition & materialize query tables.

❖ How To Implement Physical Data In Data Warehouse

- Implement physical data model transform the physical data into a physical database by generating the SQL DDL (Data Definition Language) script to create all the objects in the data.
- You implement physical data model in a production environment & populate it with data the ability to change the implementation is limited because of the data volume in a data warehouse.
- Physical data warehouse design is good query performance & this is achieved by proper implementation across all the data in the databases.
- Physical data model describe data specific implementation of the data model.
- It extracts the database & helps to generate schema.
- Physical data model implementation depends on the hardware & software used by the companies.
- Physical data can implement objects such as tables & columns are created based on entities & attributes.
- Physical data model describes data need for a single project or application through it may be integrated with other physical data models based on project scope.

- Physical data develop for a specific version of a DBMS, location, data storage or technology to be used in project.
- Physical data model describe the database specific implementation of the data model.
- The physical data model includes all required tables, columns, relationships, database properties for the physical implementation of database.
- Database performance, index strategy physical storage & de-normalization are important parameter of a physical structure.