Subject: Data Warehousing

Question no.1 Describe the Architecture of Data Warehouse?

Ans:Definition: Data warehouse architecture refers to the design of an organization's data collection and storage framework. Because data needs to be sorted, cleaned, and properly organized to be useful, data warehouse architecture focuses on finding the most efficient method of taking information from a raw set and placing it into an easily digestible structure that provides valuable

When building an organization's data warehouse, there are three main types of architecture considered, each with its own benefits and drawbacks.

- Single-Tier Architecture

- Two-Tier Architecture

- Three-Tier Architecture

  - Bottom-Tier

  - Middle-Tier

  - Top-Tier

Single tier warehouse architecture focuses on creating a compact data set and minimizing the amount of data stored. While it is useful for removing redundancies, it isn't effective for organizations with large data needs and multiple streams.

Two-tier warehouse structures separate the resources physically available from the warehouse itself. While it's more effective at storing and sorting data, it's not scalable, and it supports a minimal number of end-users.

Three tier architecture, the most popular type of data warehouse architecture, creates a more structured flow for data from raw sets to actionable insights.

The bottom tier is the database server itself and houses the back-end tools used to clean and transform data. The second tier uses Online Analytical Processing and is the go-between end-users and the warehouse. OLAPS can interact with both relational and multidimensional databases, which lets them collect data better based on broader parameters. The top tier is the front-end of an organization's overall business intelligence suite. This is where users can interact with data via queries, Data visualization, and data analytic tools.

Questions no.2 Describe star schema with example of any relevant database structure and its representation?

Ans: star schema is the basic and simplest form of a dimensional model in which data is organized into facts, dimensions and table. Star schema is use to create or develop data warehouses. Star schema is used for optimizing of query large data sets. Sales price, quantity, distant, speed, weight are the few example of fact data in star schema. The center of star can have one fact table and a number of associated dimension tables. Its structure resembles like a star.

Star schema example: sales: the sale report is one today's most common reports. In most cases we could generate sales report from the live system. But we have to build a data warehouse to streamline the process. After designing our star schema, an ETL process will get the data from operational database, transform the data in to proper format for DWH and load the data in to the warehouse.

Question no.3 Describe snowflake schema with example of any relevant database structure and its representation.?

Ans: the snowflake schema is a centralized fact tables which are connected to multiple dimensions. Basically snowflake schema is a extension of a star schema, and it adds additional dimensions. The reason why it's called snowflake schema because of its diagrammatical shape. The dimension table are normalized which splits data into additional tables.

Example: snowflake schema in sales models. This would be akin to a data mart used to track sales and results. This model has five dimensions: like product, time, sales type and employee. The fact sales table, price, and quantity are stored and grouped based on values in dimension tables. The groups in this model are same as star schema.