ADBMS UT-2.

Page No.

Q1] (1) Explain the term data warehouse

→ (i) A data warehouse is a subject oriented, integrated, non-volatile, and time-variant collection of data in support of management's decision.

(ii) A data warehouse is a type of computer database that is responsible for collecting and storing the information of a particular organization.

(4) (a) Differentiate b/w ~~operational database~~ OLAP and ~~decision support system~~ OLTP.

→

OLTP ~~operational database~~	OLAP ~~Decision support system~~
(i) OLTP ~~System~~ focus on individual database transaction.	(i) OLAP ~~System~~ focus is on how executive users like managers view the business.
(ii) An OLTP ~~System~~ maintains micro-level transaction.	(ii) This shows business trends and operations.
(iii) Data at detail level is necessary to run the business	(iii) Information concentrates on a business process.
(iv) This gives an atomic detailed data.	(iv) This ~~gives~~ gives a summarized data.
(v) The complexity of query in OD is simple to medium	(v) The ~~complexity~~ query processing is ~~extremely~~ very complex.

(3) State advantages of data warehouse.

→ (I) Enhances performance.

(i) As ~~query~~ query execution does not involve data translation with remote ~~sources~~ sources, complex queries can be executed easily.

(II) Uniformity.
(i) End users can use a single data model and query language.
(III) Potential high returns on investments
(i) Users will be able to access large amount of data.
(IV) Secure Information.
(i) Information at the warehouse is under the control of knowledged users.
(V) Competitive advantage.
(i) This can be achieved by adding DSS to data warehouse that can give us previously unavailable information.

(5) (6). Diff b/w Data warehouse and data mart.

Data warehouse.	Data Mart.
(i) Corporate/Enterprise wide data.	(i) Departmental wise data.
(ii) Aggregation of all data marts	(ii) A Single business process.
(iii) Data recieved from staging area.	(iii) Data received from facts and dimensions.
(iv) Queries on presentation resource.	(iv) Data access and analysis is simple
(v) Structure to suit for corporate view of data.	(v) Structure to suit the departmental view of data.

(6) Write a note on metadata.

→ (i) Metadata in a data warehouse is similar to the data dictionary in a DBMS, where we can keep the information about the data structures and information abt files, folders and address.

(ii) Metadata component is the data about the data in data warehouse, Metadata in a DW is much more than a data dictionary.

(iii) Types of Metadata.

(I) Operational Metadata

(II) Extraction and transformation Metadata.

(III) End-User Metadata.

(8) State the types of data warehouse (DW).

→ (i) Host based data warehouse

(ii) Host based single stage DW.

(iii) LAN based workgroup DW

(iv) Multistage DW.

(v) Stationary DW.

(vi) Distributed DW.

(vii) Virtual DW.

(9) Disadvantages of DW.

→ (a) Complexity of Integration.

(i) We need to spend lot of time on, how we are going to integrate multiple data warehousing tool.

(b) Time consuming process.

(i) Data warehouse is huge collection of data hence for each access or for any data warehouse operation we need large amt of time.

(c) Changing requirements for end-user.

(d) High investments on initial setup.

(e) High maintenance cost.

(10) State and explain operations on file ?

→ (a) OPEN:- Reads the files for access.
(b) FIND:- Searches for the first file record that satisfies a given condition and makes that record as current file record.
(c) READ:- Reads the current file record into the file.
(d) INSERT:- Adds a new record in the file.
(e) DELETE:- Removes the current record from file.
(f) MODIFY:- Changes value of some fields in current file record.
(g) CLOSE:- Terminates access to the file.

8.6 Data Warehouse Architecture

Q.2.1

MU - Dec. 2014

1. Introduction

- Data warehouse architecture is primarily based on the business processes of a business enterprise taking into consideration the data consolidation across the business enterprise with adequate security, data modeling and organization, extent of query requirements, Meta data management and application, warehouse staging area planning for optimum bandwidth utilization and full technology implementation.

2. Main areas of data warehouse

a) Data acquisition
b) Data storage
c) Information delivery

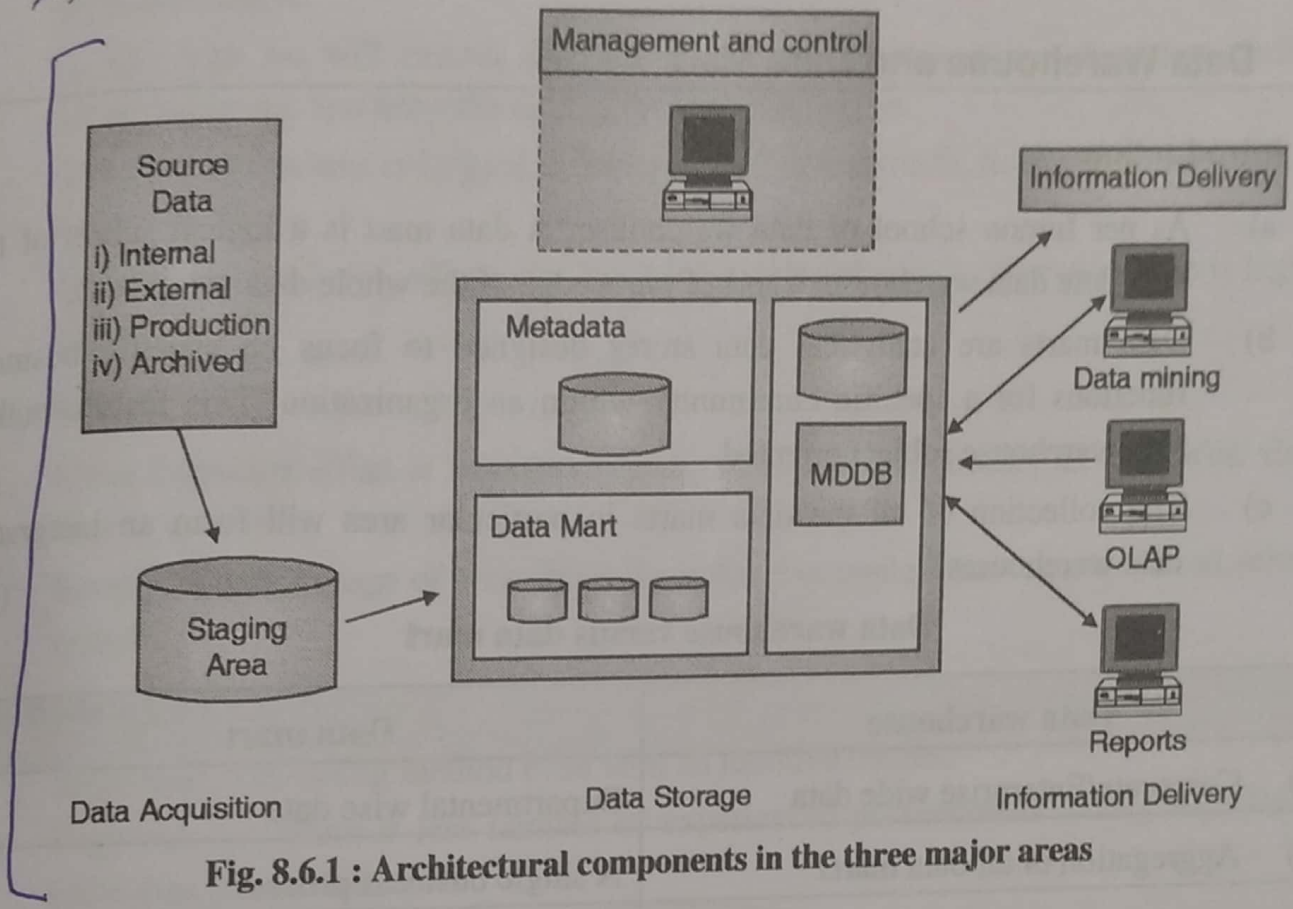

Fig. 8.6.1 : Architectural components in the three major areas

3. Building blocks of the data warehouse

a) Source data
b) Data staging
c) Data storage
d) Information delivery
e) Metadata
f) Management and control

4. Data warehouse Architecture

a) In order to set up this information delivery system, we need different building blocks. These building blocks are arranged together in the most optimal way to serve intended target.

b) Architecture, in the context of an organization's data warehousing efforts, is a conceptualization of how the data warehouse is built.

c) Data warehouse relates all components (which has definite functions and provides specific services together) so as to make fully functional data warehouse.

d) Architecture is the proper arrangement of the components.

e) We can build a data warehouse with software and hardware components.

(b) Need of data warehouse in real world - examples

1. Problem

Samtak Ltd is a company with branches at Mumbai, Delhi, Goa and Pune. The Sales Manager wants quarterly sales report for each branch but each branch has a separate operational system.

- Solution Given By Data Warehouse
  - Extract sales information from each operational database of each site.
  - Store the information in a common repository (Data warehouse) at a single site.
  - From common repository which has historic data of all three sites we can produce quarterly sales report for sales manager.

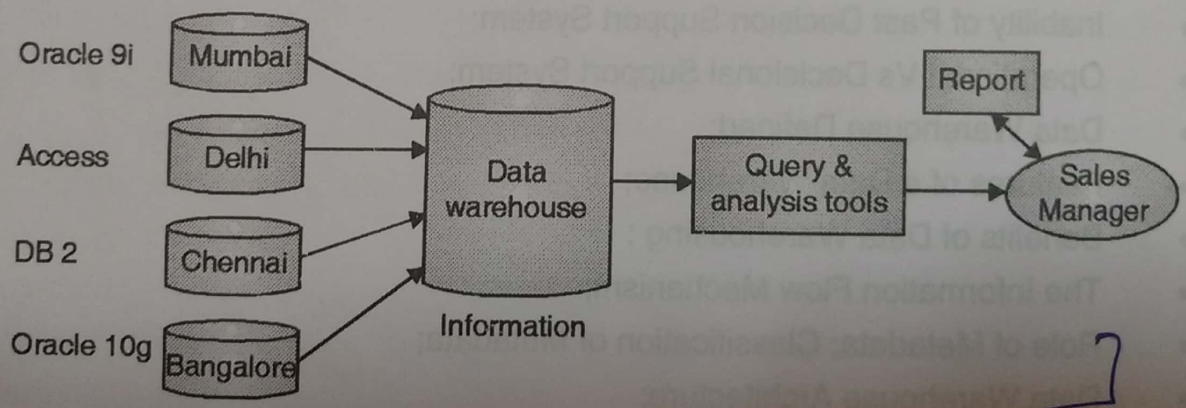

Fig. 8.1.1 : Solution for Problem (i)

Problem

Mastek Ltd. is a new company. President of the company wants his company should grow. He needs information so that he can make correct decisions.

- Solution Given By Data Warehouse
  - Extract data needed for analysis from operational database (Existing database systems).
  - Improve the quality of data before loading it into the common repository (Data warehouse) by filtering it.
  - Perform data cleaning and transformation before loading the data.
  - Use query analysis tools to support adhoc queries for taking out information for decision makers (President of company).