Data Warehouse Storage

We can divide IT systems into transactional (OLTP) and analytical (OLAP). In general we can assume that OLTP systems provide source data to data warehouses, whereas OLAP systems help to analyze it.

OLTP (On-line Transaction Processing) is characterized by a large number of short on-line transactions (INSERT, UPDATE, DELETE). The main emphasis for OLTP systems is put on very fast query processing, maintaining data integrity in multi-access environments and an effectiveness measured by number of transactions per second. In OLTP database there is detailed and current data, and schema used to store transactional databases is the entity model (usually 3NF).
Online transactional processing (OLTP) is designed to efficiently process high volumes of transactions, instantly recording business events (such as a sales invoice payment) and reflecting changes as they occur.

OLAP (On-line Analytical Processing) is characterized by relatively low volume of transactions. Queries are often very complex and involve aggregations. For OLAP systems a response time is an effectiveness measure. OLAP applications are widely used by Data Mining techniques. In OLAP database there is aggregated, historical data, stored in multi-dimensional schemas (usually star schema).

The following table summarizes the major differences between OLTP and OLAP system design.

	OLTP System Online Transaction Processing (Operational System)	OLAP System Online Analytical Processing (Data Warehouse)
Source of data	Operational data; OLTPs are the original source of the data.	Consolidation data; OLAP data comes from the various OLTP Databases
Purpose of data	To control and run fundamental business tasks	To help with planning, problem solving, and decisionsupport
What the data	Reveals a snapshot of ongoing business processes	Multi-dimensional views of various kinds of business activities
Inserts and Updates	Short and fast inserts and updates initiated by end users	Periodic long-running batch jobs refresh the data
Queries	Relatively standardized and simple queries Returning relatively few records	Often complex queries involving aggregations
Processing Speed	Typically very fast	Depends on the amount of data involved;

		batch datarefreshes and complex queries may take many hours; query speed can be improved by creating indexes
Space Requirements	Can be relatively small if historical data is archived	Larger due to the existence of aggregation structures and history data; requires more indexes than OLTP
Database Design	Highly normalized with many tables	Typically de-normalized with fewer tables; use of star and/or snowflake schemas
Backup and Recovery	Backup religiously; operational data is critical to run the business, data loss is likely to entail significant monetary loss and legal liability	Instead of regular backups, some environments may consider simply reloading the OLTP data as a recovery method

Difference Between OLAP and Data Warehouse

Data warehouse and OLAP are terms which are often used interchangeably. Actually they refer to two different components of a decision support system. While data in a data warehouse is composed of the historical data of the organization stored for end user analysis, OLAP is a technology that enables a data warehouse to be used effectively for online analysis using complex analytical queries. The differences between OLAP and data warehouse is tabulated below for ease of understanding:

Data Warehouse

- Data from different data sources is stored in a relational database for end use analysis
- Data is organized in summarized, aggregated, subject oriented, non volatile patterns.
- Data in a data warehouse is consolidated, flexible collection of data Supports analysis of data but does not support online analysis of data.

Online Analytical Processing

- A tool to evaluate and analyze the data in the data warehouse using analytical queries.
- A tool which helps organize data in the data warehouse using multidimensional models of data aggregation and summarization.
- Supports the data analyst in real time and enables online analysis of data with speed and flexibility.

A data warehouse is a database containing data that usually represents the business history of an organization. This historical data is used for analysis that supports business decisions at many levels, from strategic planning to performance evaluation of a discrete organizational unit. Data in a data warehouse is organized to support analysis rather than to process real-time transactions as in online transaction processing systems (OLTP).

OLAP technology enables data warehouses to be used effectively for online analysis, providing rapid responses to iterative complex analytical queries. OLAP's multidimensional data model and data aggregation techniques organize and summarize large amounts of data so it can be evaluated quickly using online analysis and graphical tools. OLAP systems provide the speed and flexibility to support the analyst in real time.

More   differences:

Sr.No.	Data Warehouse (OLAP)	Operational Database(OLTP)
1	It involves historical processing of information.	It involves day-to-day processing.
2	OLAP systems are used by knowledge workers such as executives, managers, and analysts.	OLTP systems are used by clerks, DBAs, or database professionals.
3	It is used to analyze the business.	It is used to run the business.
4	It focuses on Information out.	It focuses on Data in.
5	It is based on Star Schema, Snowflake Schema, and Fact Constellation Schema.	It is based on Entity Relationship Model.
6	It focuses on Information out.	It is application oriented.
7	It contains historical data.	It contains current data.
8	It provides summarized and consolidated data.	It provides primitive and highly detailed data.
9	It provides summarized and multidimensional view of data.	It provides detailed and flat relational view of data.
10	The number of users is in hundreds.	The number of users is in thousands.
11	The number of records accessed is in millions.	The number of records accessed is in tens.
12	The database size is from 100GB to 100 TB.	The database size is from 100 MB to 100 GB.

13	These are highly flexible.	It provides high performance.