|  |  |
| --- | --- |
| OLAP | OLTP |
| * Uses information in database to guide strategic decisions   + Complex queries   + Infrequent Updates   + Large transactions that would take a lot of computer power   + More historic data | * Maintains a Database that is accurate model of real world enterprise * Constructed to support **day to day tasks** and **constant update on DBMS**   + Short **simple transactions**   + Relatively **frequent updates**   + Only accesses a **fraction of the databases** |
| Complex SQL Queries and Views in order to **create and produce an analytical report** based on the data and facts |  |
| Updated **periodically**   * Updating OLAP takes a large sum of computer power | Updated **Every time** |
| When you Aggregate or analyses from the data | When you query or add/insert/update to the data |

Data Warehouse

* Special server where OLAP and Data Minding operations are done and stored
  + Subject oriented, integrated, time variant, non-updatable collections of data
  + Can accommodate **large sets of data**
  + **Allows OLAP and Data Mining** to be operated **independently** from main **OLAP server**

## Issues with Data Warehouse

|  |  |
| --- | --- |
| Semantic Integration | When getting data from **multiple sources,** must eliminate **mismatching data** or **accommodate different data integrity** |
| Heterogenous Sources | Must **access data** from variety of **source format and repository** |
| Load, refresh, purge | Must load data, periodically refresh and purge old data   * Due to the volume of data, these takes take up a lot of computer power   + Data integrity checking and consistency translations |
| Metadata management | Must keep track of all the **metadata information** about **all the data in the warehouse** |

# Data Lake

* Enterprises getting a lot of cheap CPU and Storage unit to store **old historic data**.
  + This is because getting **ONE MASSIVE COMPUTER** will cost them more than buying **several small cheap ones**

|  |  |
| --- | --- |
| Relational OLAP | Multi-Dimensional OLAP |
| * Data is stored in relational database | * One **BIG ASS TABLE** that holds all the informations * NON-RELATIONAL |

## Fact Table (Multi-dimensional Table for Multi-dimensional OLAP)

* Table that covers all the tables relevant to the Business Process
* Collection of numeric measures, which if

## Star Schema

* The fact and dimensions relations can be displayed in an ER diagram which has **fact table at the center** with radiating entities as **different dimensions** that hold **primary-foreign key relationship**
  + 1 Central **fact table**
  + N Dimensions table with **foreign key** relationships **from the fact table**

