2.3 CHARACTERISTICS OF OLAP SYSTEMS

The following are the differences between OLAP and OLTP systems.

1. **Users**: OLTP systems are designed for office workers while the OLAP systems are designed for Decision makers. Therefore while an OLTP system may be accessed by hundreds or even thousands of users in a large enterprise, an OLAP system is likely to be accessed only by a select group of managers and may be used only by dozens of users.

2. **Functions**: OLTP systems are mission-critical. They support day-to-day operations of an enterprise and are mostly performance and availability driven. These systems carry out simple repetitive operations. OLAP systems are management-critical to support decision of an enterprise support functions using analytical investigations. They are more functionality driven. These are ad hoc and often much more complex operations.

3. **Nature**: Although SQL queries often return a set of records, OLTP systems are designed to process one record at a time, for example a record related to the customer who might be on the phone or in the store. OLAP systems are not designed to deal with individual customer records. Instead they involve queries that deal with many records at a time and provide summary or aggregate data to a manager. OLAP applications involve data stored in a data warehouse that has been extracted from many tables and perhaps from more than one enterprise database.

4. **Design**: OLTP database systems are designed to be application-oriented while OLAP systems are designed to be subject-oriented. OLTP systems view the enterprise data as a collection of tables (perhaps based on an entity-relationship model). OLAP systems view enterprise information as multidimensional).

5. **Data**: OLTP systems normally deal only with the current status of information. For example, information about an employee who left three years ago may not be available on the Human Resources System. The old information may have been achieved on some type of stable storage media and may not be accessible online. On the other hand, OLAP systems require historical data over several years since trends are often important in decision making.

6. **Kind of use**: OLTP systems are used for reading and writing operations while OLAP systems normally do not update the data.

The differences between OLTP and OLAP systems are:
### Table 2.2: Comparison of OLTP and OLAP Systems

<table>
<thead>
<tr>
<th>Property</th>
<th>OLTP</th>
<th>OLAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature of users</td>
<td>Operations workers</td>
<td>Decision makers</td>
</tr>
<tr>
<td>Functions</td>
<td>Mission-critical</td>
<td>Management-critical</td>
</tr>
<tr>
<td>Nature of queries</td>
<td>Mostly simple</td>
<td>Mostly complex</td>
</tr>
<tr>
<td>Nature of usage</td>
<td>Mostly repetitive</td>
<td>Mostly ad hoc</td>
</tr>
<tr>
<td>Nature of design</td>
<td>Application oriented</td>
<td>Subject oriented</td>
</tr>
<tr>
<td>Number of users</td>
<td>Thousands</td>
<td>Dozens</td>
</tr>
<tr>
<td>Nature of data</td>
<td>Current, detailed, relational</td>
<td>Historical, summarized, multidimensional</td>
</tr>
<tr>
<td>Updates</td>
<td>All the time</td>
<td>Usually not allowed</td>
</tr>
</tbody>
</table>

**FASMI Characteristics**

In the FASMI characteristics of OLAP systems, the name derived from the first letters of the characteristics are:

**Fast:** As noted earlier, most OLAP queries should be answered very quickly, perhaps within seconds. The performance of an OLAP system has to be like that of a search engine. If the response takes more than say 20 seconds, the user is likely to move away to something else assuming there is a problem with the query. Achieving such performance is difficult. The data structures must be efficient. The hardware must be powerful enough for the amount of data and the number of users. Full pre-computation of aggregates helps but is often not practical due to the large number of aggregates. One approach is to pre-compute the most commonly queried aggregates and compute the remaining on-the-fly.

**Analytic:** An OLAP system must provide rich analytic functionality and it is expected that most OLAP queries can be answered without any programming. The system should be able to cope with any relevant queries for the application and the user. Often the analysis will be using the vendor’s own tools although OLAP software capabilities differ widely between products in the market.

**Shared:** An OLAP system is shared resource although it is unlikely to be shared by hundreds of users. An OLAP system is likely to be accessed only by a select group of managers and may be used merely by dozens of users. Being a shared system, an OLAP system should be provide adequate security for confidentiality as well as integrity.
**Multidimensional**: This is the basic requirement. Whatever OLAP software is being used, it must provide a multidimensional conceptual view of the data. It is because of the multidimensional view of data that we often refer to the data as a cube. A dimension often has hierarchies that show parent/child relationships between the members of a dimension. The multidimensional structure should allow such hierarchies.

**Information**: OLAP systems usually obtain information from a data warehouse. The system should be able to handle a large amount of input data. The capacity of an OLAP system to handle information and its integration with the data warehouse may be critical.

**Codd’s OLAP Characteristics**

Codd et al’s 1993 paper listed 12 characteristics (or rules) OLAP systems. Another six in 1995 followed these. Codd restructured the 18 rules into four groups. These rules provide another point of view on what constitutes an OLAP system. All the 18 rules are available at [http://www.olapreport.com/fasmi.htm](http://www.olapreport.com/fasmi.htm). Here we discuss 10 characteristics that are most important.

1. **Multidimensional conceptual view**: As noted above, this is central characteristic of an OLAP system. By requiring a multidimensional view, it is possible to carry out operations like slice and dice.

2. **Accessibility (OLAP as a mediator)**: The OLAP software should be sitting between data sources (e.g. data warehouse) and an OLAP front-end.

3. **Batch extraction vs. interpretive**: An OLAP system should provide multidimensional data staging plus pre calculation of aggregates in large multidimensional databases.

4. **Multi-user support**: Since the OLAP system is shared, the OLAP software should provide many Normal database operations including retrieval, update, concurrency control, integrity and security.

5. **Storing OLAP results**: OLAP results data should be kept separate from source data. Read-write OLAP applications should not be implemented directly on live transaction data if OLAP source systems are supplying information to the OLAP system directly.
6. **Extraction of missing values:** The OLAP system should distinguish missing values from zero values. A large data cube may have a large number of zeros as well as some missing values. If a distinction is not made between zero values and missing values, the aggregates are likely to be computed incorrectly.

7. **Treatment of missing values:** An OLAP system should ignore all missing values regardless of their source. Correct aggregate values will be computed once the missing values are ignored.

8. **Uniform reporting performance:** Increasing the number of dimensions or database size should not significantly degrade the reporting performance of the OLAP system. This is a good objective although it may be difficult to achieve in practice.

9. **Generic dimensionality:** An OLAP system should treat each dimension as equivalent in both is structure and operational capabilities. Additional operational capabilities may be granted to selected dimensions but such additional functions should be grantable to any dimension.

10. **Unlimited dimensions and aggregation levels:** An OLAP system should allow unlimited dimensions and aggregation levels. In practice, the number of dimensions is rarely more than 10 and the number of hierarchies rarely more than six.