ETL Process

Over the course of this week. I have come to learn more about the ETL process and why we need an ETL. Also, I have learned more about the bridge an ETL can make between an OLTP and an OLAP. During the process of an ETL, I have learned about what happens in the following stages. First, the extract stage, second, the transformation stage, and third, the load stage. Lastly, I learned more about the role of data integration in the ETL process, the role of programming languages in the ETL process, and the role of the command-line tools in the ETL process.

The reason a company would want an ETL is because an ETL enhances business intelligence. It enables companies a consolidated data view to drive better and more informed decisions1. The bridge that an ETL can make between an OLTP and an OLAP is a great value to the company. An ETL process will change the data OLTP into data OLAP. With this, the ETL bridges the gap between the two.

There are three main stages within an ETL process. These stages consist of the following, the extract stage, the transformation stage, and the load stage. In the extraction stage, data is extracted from the source system into the staging area. Within the transformation stage, the data is extracted from the source is cleansed and transformed. Lastly, in the loading stage, the data is loaded into the target data warehouse.

Data integration within the ETL process was done primarily by ETL developers. They used hand coding or specialized software to allow data to move from one machine to the next. Programming languages are what create the ETL to allow the data to flow from machine to machine. Lastly, the role of command-line tools help to automate the extraction, transforming, and loading processes of an ETL.

With so much learned and even more to be learned. The ETL process is key to allowing the flow of data from one machine to another. This includes the programmers and the data scientist. This includes the bridges that ETL’s can create for OLTP and OLAP systems. ELTs are what allows the data to flow properly.

References

1. <https://datafortune.com/what-are-the-benefits-of-etl-for-your-company/>
2. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://iopscience.iop.org/article/10.1088/1757-899X/128/1/012020/pdf
3. <https://www.guru99.com/etl-extract-load-process.html#:~:text=In%20the%20first%20step%20extraction,step%20of%20the%20ETL%20process>.
4. <https://streamsets.com/learn/data-integration/?gclid=Cj0KCQiA8aOeBhCWARIsANRFrQGhhWb9KPiIDbLGsA2VQ4UczXf6cZ-Xg2z4kDRQWrPlvKMFeEnwyaMaAl46EALw_wcB>
5. <https://www.altexsoft.com/blog/datascience/who-is-etl-developer-role-description-process-breakdown-responsibilities-and-skills/>
6. <https://www.snowflake.com/guides/what-etl>