REAL-TIME ETLT: MEETING THE DEMANDS OF MODERN DATA PROCESSING

A businessman had two lathes in his factory, one manual, and the other computer numerically controlled (CNC). The manual lathe turned metal quite efficiently, but at a slower rate, while the CNC lathe turned three times as much as the manual lathe. All the CNC lathe needed was accurate data entry.

Modern data processing is like a CNC lathe, helping us manage large amounts of data effectively and efficiently. Real-time ETLT can help us achieve our goals in today's fast-paced business environment. They provide invaluable insight into market trends by focusing on the most important asset, DATA.

INTRODUCTION

A deep look at what ETLT stands for

- Data Extract (E) - Data is extracted from various sources.
- Data Transform (T) - The extracted data is transformed into a single unified format by filtering the data.
- Data Load (L) - The transformed data is loaded into the target database.
- Data Transfer (T) - The loaded data is then transferred to a repository.

Real-ETLT enables modern organizations to respond quickly to emerging trends, market changes, and new opportunities in real-time. Organizations must ensure data quality, security, and consistency, especially when handling sensitive customer information. For longevity, organizations must cultivate a compliant culture of data on their business objectives

CHALLENGES OF REAL-TIME ETLT

- Real-time ETLT manages large data sets that it documents in the database and quickly transmits them to a real-time system.
- Due to the endless flow of data, it becomes complicated to modify the old data and update the new data, which cannot guarantee the accuracy of the data.
- Real-time ETLT tools manage the high risks of cyber-attacks, data breaches, and malware; managing and maintaining this data is essential.

SOLUTIONS FOR REAL-TIME ETLT

- Stream Processing Technologies - Organizations can leverage streaming processes such as Apache, Kafka, Flink, etc. which are very useful for real-time data management and interpretation.
- Data Quality Control and Data Monitoring – This can help detect any inconsistency issues the data may have and guide the organization to resolve them quickly and ensure the data is working effectively.

- Establishment of secure data transfer protocols and encryption - Data encryption algorithms help identify and fix data problems before they become serious problems. Data transfer can also protect the organization from any hacking or malware activities.

REAL-TIME DATA EXTRACTION

To take advantage of real-time data, you need to identify the type of data and consider data extraction from different sources such as API, database, files, cloud, CRM systems, etc. An effective strategy capable of integrating with various tools in real-time and the chosen extraction method must be effective and up-to-date.

The organization uses refactored data that has been previously mined, this first mined data has been previously stored and not integrated with the rest of the refined data due to the risk of malicious configurations.

REAL-TIME DATA TRANSFORMATION AND LOADING

The information obtained from the extracted data is transformed into a single and unified format using real-time tools and data integration pipelines. After the transformation is successful, the data will be loaded into the repository. Due to the volume and velocity of the data, along with its persistence and reliability, it is a bit difficult to continue loading data to the system in real time. Organizations can rely on various tools such as Apache Ignite to process data across different systems.

CONCLUSION

Technology has advanced and the need for acute data management is high. ETLT in real time, its tools and processes have proven to be very reliable for organizations. It improves operational productivity through rapid decision-making based on real-time data and a little help from highly skilled engineers. Organizations can focus on a system that can handle high volumes of data, high-quality controls, and processing technology that ensures smooth data operation.