**Insights on Data Stream.**

* **What is data stream?**

Data stream" typically refers to a continuous flow of data generated from various sources, such as sensors, devices, applications, social media, or website interactions. Analyzing data streams in real-time or near-real-time can provide valuable insights into ongoing processes, trends, anomalies, and patterns.

* **Why is data stream needed?**

Here are some insights on data streams on analytics and reporting perspectives:

1.Data streams enable organizations to make faster and more informed decisions by providing real-time insights into operations, customer behavior, and market trends.

2.Analyzing historical data streams can help organizations identify patterns and trends, enabling the development of predictive models for forecasting future events or outcomes.

3.Data streams are often used for anomaly detection, where deviations from expected patterns or behaviors are flagged for further investigation.

4.Analyzing IoT data streams enables organizations to gain insights into device performance, environmental conditions, and user behavior, facilitating predictive maintenance, resource optimization, and personalized services.

5.Social media platforms generate continuous streams of data, including user-generated content, interactions, and engagement metrics.

6.Data streams play a crucial role in financial markets, where real-time data on stock prices, trading volumes, and market trends are continuously monitored and analyzed. High-frequency trading (HFT) firms leverage data streams to execute rapid trades based on algorithmic models, exploiting small price discrepancies for profit.

7.Data streams from network traffic provide insights into network performance, security threats, and user behavior. By analyzing network data streams in real-time, organizations can detect and mitigate cybersecurity threats, optimize network infrastructure, and ensure compliance with regulatory requirements.

* **How does it work?**

Data streaming is focused on capturing events at a high level. The architecture is designed such that:

1.Data is captured continuously in real-time by producers which can be databases, software application etc.

2.Data Information is sent into the stream where is it written to a log file and placed in a queue. The queue can be broken down into groups or subsections instead of a single queue for faster parallel processing.

3.Data from the queue can be streamed by consumers in near real-time which can be databases, analytical tool and other stream.

**NB**. These queues are stored in multiple serves to protect the architecture should something happen, it doesn’t impact the whole architecture.

* **What are some examples?**

Stream can be open-source service offered by several cloud providers and it can be configured to serve every individual or organisational use as it is scalable and full tolerant.

1.AWS Azure (Azure stream analytics)

2.Google cloud

3.IBM streaming analytics.

4.GCP Dataflow

* **Importance of data stream?**

1.Data streaming could capture every tracking update as it occurs in real-time, giving full visibility access.

2.Allows data to be passed between applications in real-time so that they are all on the same page.

3.It’s a more scalable solution for the massive amount of data generated every minute of every day. Moving data along as soon as it’s created avoids overloading a batch load server.

**Cons:**

The amount of data generated through stream is massive and requires a lot of technical maintenance which can be overwhelming. If the business does not require so much of real time data, using services offered by one of the cloud providers is a better solution.