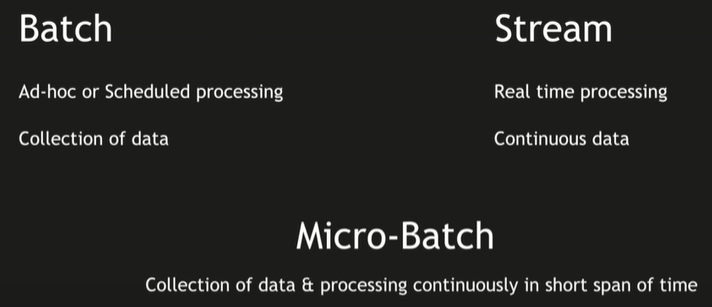
**Spring Batch:**

**Batch vs Stream:**

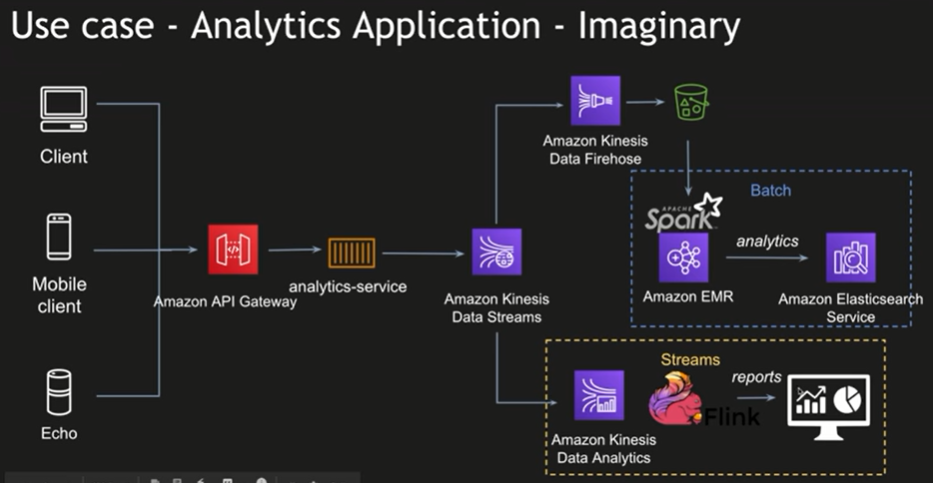


**Batch processing use case:**

* When we have scheduled processes or EOD processes
* When we have huge data/ huge files to process
* We can use big data frameworks such as Spark/Hadoop to process much faster and load into database.
* Later on we can use data stored in database, to view data on UI or for analytics purpose.

**Stream processing use case:**

* When we need to do real time processing
* We can use Kafka, where different applications can post data and we can use stream processing frameworks such as Storm, Flink, Spark, etc., to process/format data and store data in database. Also, we can post processed data which needs to be actioned upon another kafka topic.



[Netflix & Amazon Kinesis Streams Case Study](https://aws.amazon.com/solutions/case-studies/netflix-kinesis-data-streams/)

[Nasdaq’s Architecture using Amazon EMR and Amazon S3 for Ad Hoc Access to a Massive Data Set | AWS Big Data Blog](https://aws.amazon.com/blogs/big-data/nasdaqs-architecture-using-amazon-emr-and-amazon-s3-for-ad-hoc-access-to-a-massive-data-set/)

**Spring Batch Architecture:**

