Kafka / Kinesis has a good integration with Spark.

We can set the window duration to consume from various streams.

Using Kafka we can store the data for a longer duration when we need historical data.

AWS S3 / HDFS for data lake makes a perfect storage area for accessing data.

Preferable AWS S3 for its availability and scalability on-demand.

Spark for In-Memory processing is a better option to deliver the metrics computation be it near real time or a batch wise processing.