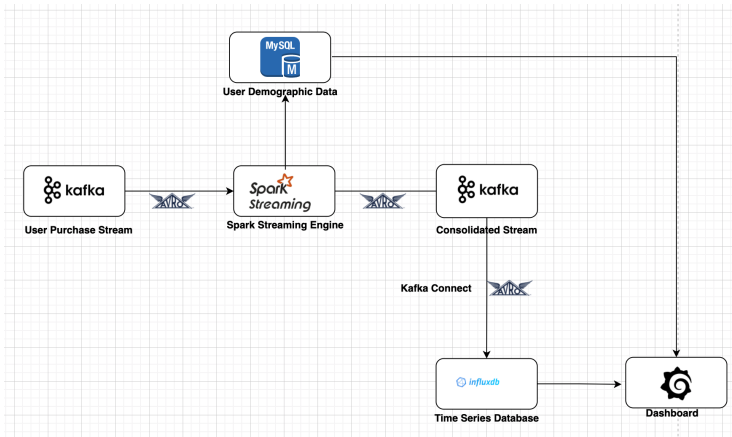


High Level Design



- User purchase events in avro format produced via kafka.
- Spark Streaming Framework does join batch and realtime events of user purchase and demographic type.
- MySql Holds the demographic data such as age, gender, country, etc
- Spark Streaming Framework consumes these events and generated different variety of events suitable for time series and dashboarding.

- Kafka connect pushes the events from the kafka streams to influxdb.
- Grafana connects to different sources like influxdb, mysql and populates the graphs.

Few Points to note :

- The batch data (demographic) is created in mysql.
- The streams of data is produced by running the java classes as mentioned in the section execution.