Click-Stream Analysis for Online Retail Stores

Team:

- ➢ Mayank Lamba (ml5711)
- ➢ Alisha Goyal (ag6599)
- ➢ Lasya Lamineni (ln1026)
- ➢ Varun Sanjay Mishra (vm1420)

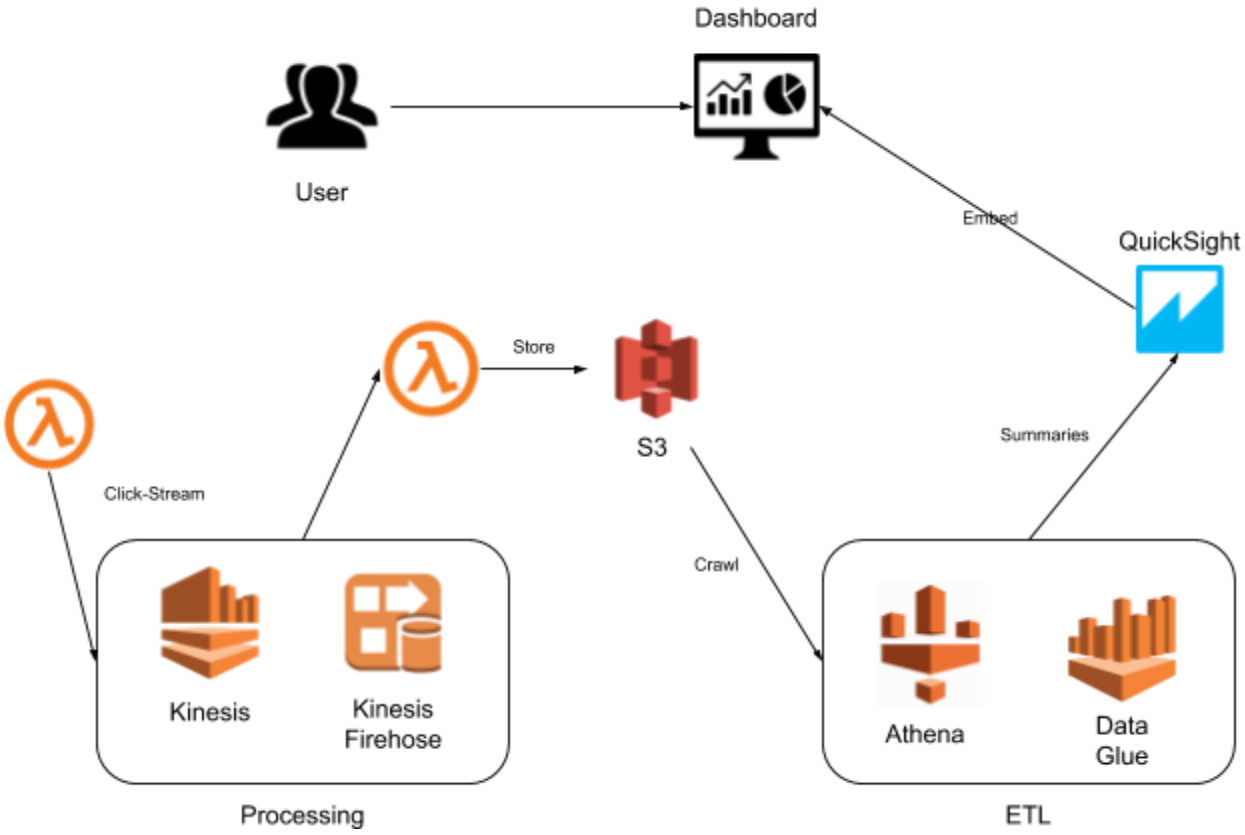
Problem Statement:
Leverage AWS technologies to build a serverless system that can **ingest, process, store and analyze** incoming stream data from an online retail store.

Overview:
The system that we are proposing to build will be targeted towards people who are trying to manage their online retail stores. We hope to provide them with an architecture which they can use to make sense of the large amount of online traffic data generated from various clicks on their website. We wish to sessionize the incoming user data and perform some basic ETL before we do the analysis. This will help us process the data in a way that we can provide actionable user insights with respect to their surfing patterns during the time they surf an online retail store website. It will help empower the store managers to stay ahead in terms of in terms of the needs of their customers. Using the AWS technologies will give them serverless architecture that can scale up and down depending upon the stream generated by their online traffic.

Implementation/Deliverables:
Following are the deliverables that we hope to implement in our web app:

- User Authentication: User (in this case the manager) has to authenticate to use the platform. We will use AWS Cognito for this purpose.
- Web Application Interface: This is where we will embed the Amazon QuickSight dashboard that will house the final visualizations.
- Data Ingestion & Sessionization: We will use Amazon Kinesis stream to get the data generated from a lambda function (essentially augmenting the click-stream data) and Kinesis Analytics to sessionize it.
- Data Processing & Storage: We will use a Lambda function to get the sessionized data and use Kinesis Firehose to store it in an S3 bucket.
- Data Analysis & ETL: AWS Data Glue will be used to crawl data from S3 and update the Amazon Athena tables.
- Visualizations: The client facing dashboard will be built using Amazon QuickSight.

Architecture:



Technology Stack:

- Cloud Tools:
  - AWS Lambda
  - AWS Cognito
  - Amazon QuickSight
  - Amazon Athena
  - AWS Data Glue
  - AWS Kinesis (Firehose, Analytics)
  - AWS S3
  - AWS API Gateway
  - AWS CloudWatchEvents
- Backend: AWS Lambda
- Frontend: HTML/CSS/JavaScript