Website Monitoring Project Overview

What is Website Monitoring?

Website Monitoring is defined as the term used for any activity which involves testing a website or any web service for the availability, performance, or the function. It's the process of testing and also verifying that the end-users can interact with the website or the web application as expected. The Website Monitoring service checks and verifies the website is up and working as expected and website visitors can use the site without facing any difficulties as expected.

Data Pipeline:

It refers to a system for moving data from one system to another. The data may or may not be transformed, and it may be processed in real-time (or streaming) instead of batches. Right from extracting or capturing data using various tools, storing raw data, cleaning, validating data, transforming data into query worthy format, visualisation of KPIs including Orchestration of the above process is data pipeline.

What is the Agenda of the project?

The agenda of the project involves Real-time monitoring of website using AWS services. We first launch an EC2 instance on AWS and install Amazon Kinesis in it. Then, Data analytics streams are created in amazon Kinesis for real-time streaming of data. Then after launching AWS EC2 instances, an Identity Access Management(IAM) role is assigned to the EC2 instance. It is followed by doing Log streaming to Kinesis data streams, followed by Creation of Kinesis Analytics. It is followed by creation of Amazon Aurora MySQL followed by using Amazon SNS service and Secret Manager. Further, AWS Lambda code is run for end to end testing and Amazon DynamoDB is run with selection of Keys. It is followed by creation of AWS Lambda for loading in Amazon DynamoDB, Loading DynamoDB with order logs and finally Real-time streaming the data in the Kinesis Analytics.

Architecture:

1) Amazon EC2 acts as website backend generating server logs.

2) Kinesis Datastreams reads the server logs in real time and pushes it to Kinesis Data Analytics for doing above computation(more than 15 orders per 15 seconds).

3) Second stream in Data Analytics is created that actually notes such floods for past 1 minute and then send only messages to 2nd Data stream if the trend is noted continuously 1 min. This step is purely added to reduce number of SNS messages received in case of spike.

4) Second data stream is used to receive such alarm records and trigger lambda.

5) Lambda triggers SNS notification which in turn delivers an SMS message. It saves the copy of all the error messages in Aurora MySQL for aggregated view in future .

Key Takeaways

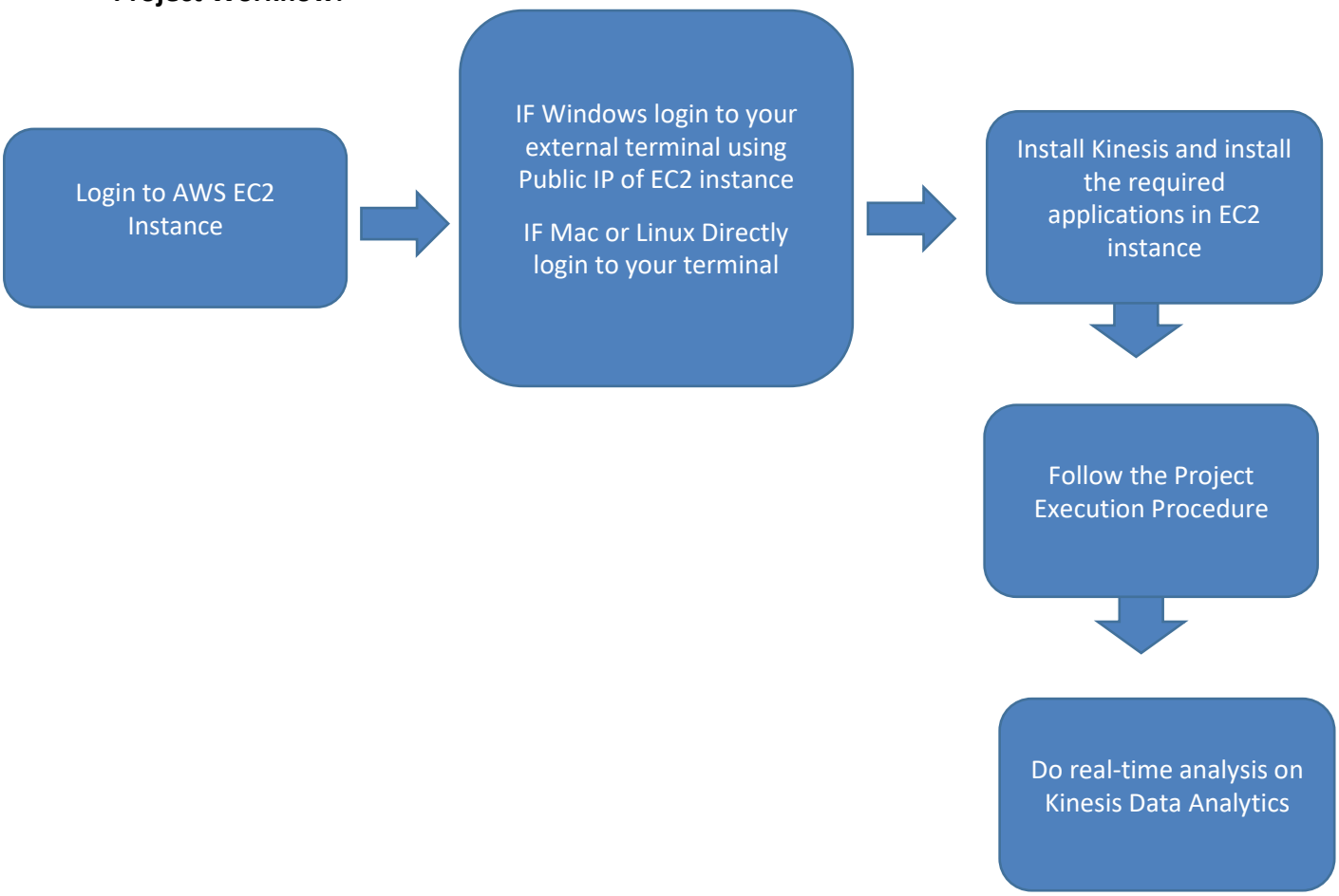
- Understanding the project and how to use AWS EC2 Instance
- Understanding the basics of Serverless computing, MySQL, and their application
- Doing real-time streaming the data in Amazon Kinesis Data Analytics application
- Installing Amazon Kinesis in EC2 instance
- Usage of Kinesis data streams and doing real-time streaming

- Exploring Amazon Kinesis by doing Data Analytics and Log streaming
- Using Amazon DynamoDB for creating NoSQL database.
- Using Amazon RDS for MySQL
- Using Amazon Aurora for MySQL
- Using Amazon SNS for simple notification service
- Integration of Amazon Kinesis data stream and Kinesis Data Analytics
- Using end to end testing of AWS lambda code
- Integration of Amazon Aurora and AWS lambda
- Selection of Keys in Amazon Dynamo DB
- Integration of Amazon SNS and AWS lamba
- Loading of Amazon DynamoDB in AWS lambda
- Loading of Amazon DynamoDB with order logs
- Displaying real-time streaming data of website using Amazon Kinesis Data Analytics application

Data Analysis:

- From the given website, PUTTY is downloaded to run Linux shell in Windows for running website monitoring using Amazon Kinesis.
- AWS EC2 instance is created and PUTTY shell is connected to EC2 instance. Amazon Kinesis is downloaded in the EC2 instance followed by addition of python files and dataset.
- The Amazon Kinesis data streams are creating in the AWS console followed by attaching Identity Access Management(IAM) to the EC2 instances.
- The Amazon Kinesis analytics is created for performing analytics of website real-time streaming. Amazon Aurora MySQL is created for using relational database.
- The AWS Lambda is created and its code is executed for end to end testing so as to enable smooth analytics.
- The Amazon DynamoDB is created for NoSQL database and keys are selected in it. Amazon DynamoDB is loaded in AWS Lambda.
- Finally, Amazon DynamoDB is loaded with order logs and data is analysed in real-time streaming in Amazon Kinesis Data Analytics application.

Project Workflow:



Folder Structure: