Real Time Data Processing Application

Submitted by:-Saksham,Varun,Avani,Pragati

Acknowledgement

We wish to express our deep sense of gratitude to our Project Coordinator **Mr. Naveen Kumar Kedia** and Project Guide **Ms. Shweta Saxena** , Jaipur Engineering College and Research Centre, Jaipur for guiding us from the inception till the completion of the project. We would like to first of all express our thanks to **Mr. Arpit Agrawal**, Director of JECRC, for providing us such a great infrastructure and environment for our overall development.

We express sincere thanks to **Dr. V. K. Chandna**, Principal of JECRC, for his kind cooperation and extendible support towards the completion of our project.

Words are inadequate in offering our thanks to **Dr. Sunil Kumar Jangir**, HOD of Department of Information Technology, for consistent encouragement and support for shaping our project in the presentable form. Also our warm thanks to **Jaipur Engineering College and Research Centre**, who provided us this opportunity to carry out this prestigious Project and enhance our learning in various technical fields.

Project Background

Cloud computing has been around for approximately two decades and despite the data pointing to the business efficiencies, cost benefits and competitive advantages it holds over the old way of business, a large portion of the business community continues to operate the old way.; According to a study by the International Data Group, 69 per cent of businesses are already using cloud technology in one capacity or another, and 18 per cent say they plan to implement cloud-computing solutions at some point.
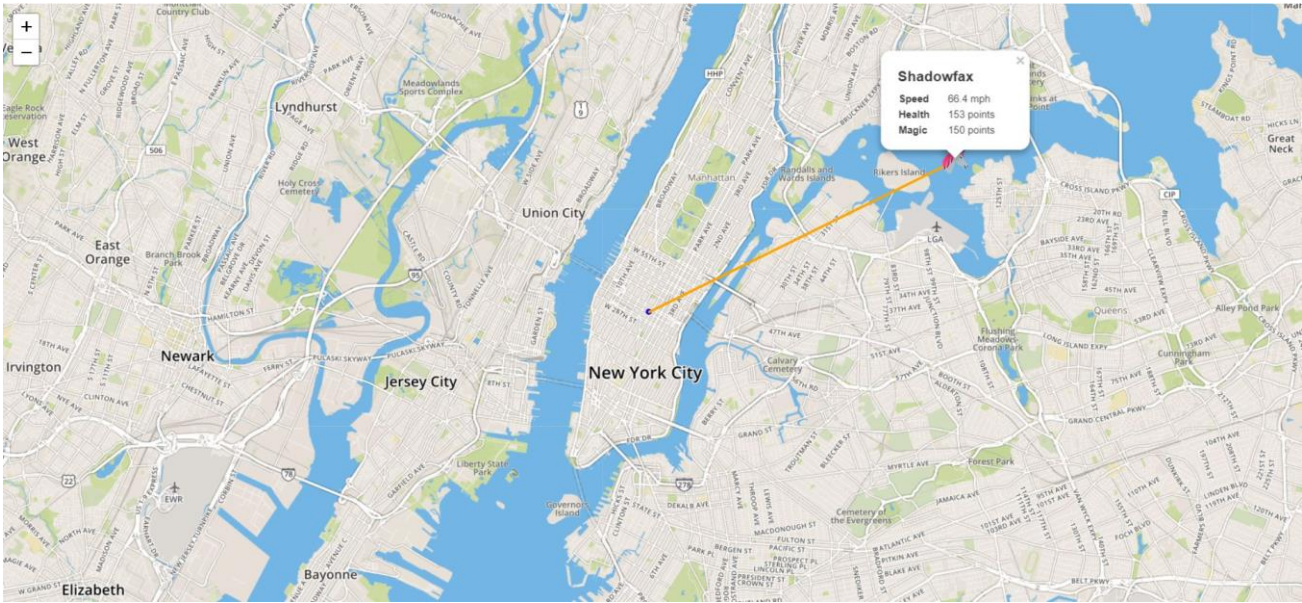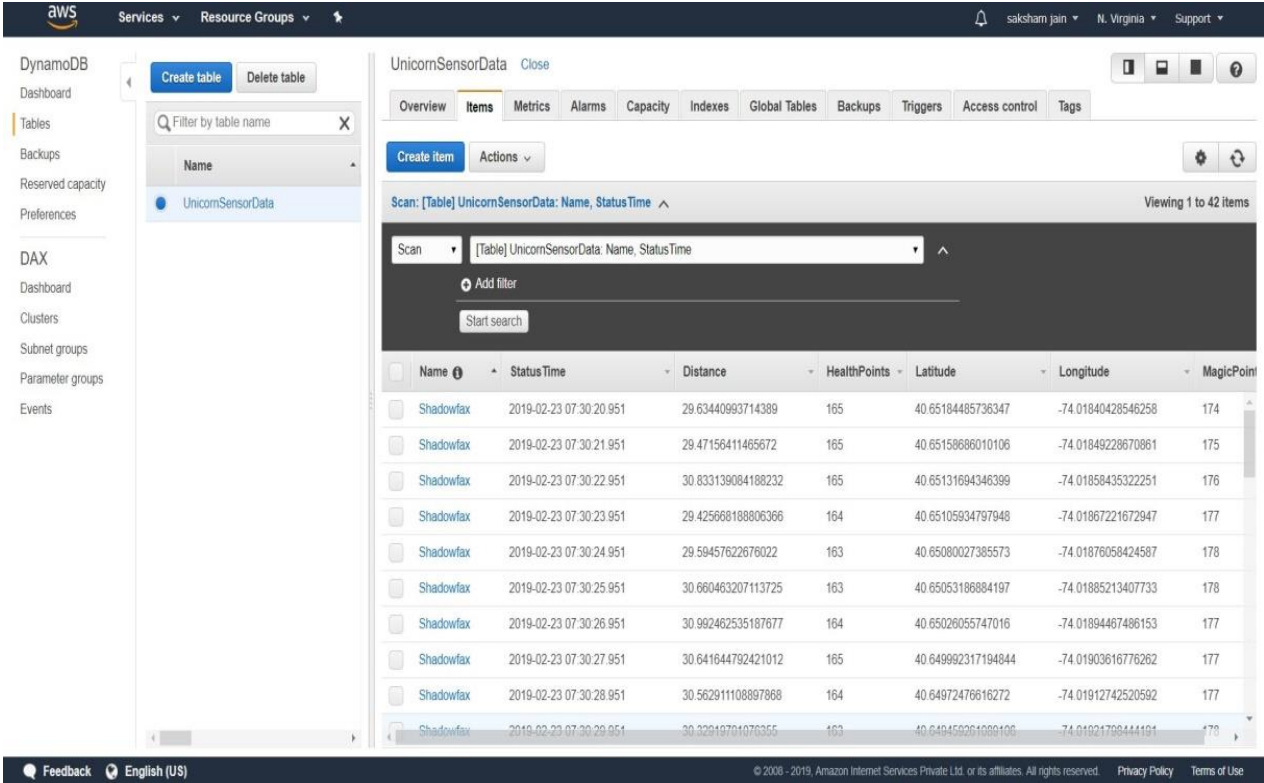
Cloud computing operates on a similar principle as web-based email clients, allowing users to access all of the features and files of the system without having to keep the bulk of that system on their own computers. In fact, most people already use a variety of cloud computing services without even realizing it's Gmail, Google Drive, TurboTax, and even Facebook and Instagram are all cloud-based applications..

Purpose

By processing the data in motion, real-time Big Data Processing enables you to walk in parallel with the current landscape of your Business and turn data intelligence into vital business decisions. Whether it is positive, negative or neutral, a clear picture can be visualized about the current status of the projects.

The demand of Real Time Streaming Platforms is high these days. No doubt that the data being generated today every second is colossal, but what use is the data of if it is not processed to draw conclusions, trends, patterns, and outliers? Not merely the processing, but a quick processing should be done quickly so that a firm can react to the changing patterns in the business in real-time.

Screenshots/Results

Methodology

We will use AWS for to build applications to process and visualize this data in real-time. We use AWS services like aws lambda function, kinesis data firehose, kinesis data analytics application, DynamoDB as a database and divide the project in 4 modules:-

1. Build a data stream-
Create a stream in Kinesis and write to and read from the stream to track Wild Rydes unicorns on the live map. In this module we'll also create an Amazon Cognito identity pool to grant live map access to our stream.

2. Aggregate data-
We will build a Kinesis Data Analytics application to read from the stream and aggregate metrics like unicorn health and distance traveled each minute and location.

3. Process streaming data-
Store aggregate data from the application to a backend database stored in DynamoDB and run queries against those data.

4. Store & query data-
We will use Kinesis Data Firehose to flush the raw sensor data to an S3 bucket for archival purposes. Using Athena, we'll run SQL queries against the raw data for ad-hoc analyses.

Tools and Technology Used

Since This Project is developed on cloud so there is need of computer with minimum hardware configurations like :-

- o Processor of Pentium or above.
- o Minimum of 256 MB RAM.
- o Minimum of 20 GB Hard disk.
- ➢ Working network (LAN)

And internet connectivity must be present for working of project on the computer.
Database

- DynamoDB
- SQL
- AWS Athena

References

REFERENCE BOOKS REFERRED:

- Learn SQL: Book by Jose A. Ramalho
- Node.js in Action: Book by Marc Harter, Mike Cantelon, Nathan Rajlich, and T. J. Holowaychuk
- Mastering DynamoDB: Book by Tanmay Deshpande
- Beginning Node.js :Book by Basarat Ali Syed

AWS deployment-

- https://docs.aws.amazon.com