**Business Problem & Scope Understanding**

**About Customer**

**Apty** is a Digital Adoption Platform that helps enterprise optimize their business processes. Successful digital adoption involves both guiding people through new critical software applications and proactively pushing them to complete new processes. Apty combines the power of on-screen guidance with the time-saving automation of proactive process compliance. IT operation managers, Application Owners, VP HR/Sales, and leading CIOs alike use Apty to get the most out of their employees' use of web-based applications in their day-to-day job. Every month, nearly 1 million users use Apty to learn web-based apps and excel in their job in leading companies like Delta Airlines, Boeing, Agile CRM, and many more.

**Business Objective**

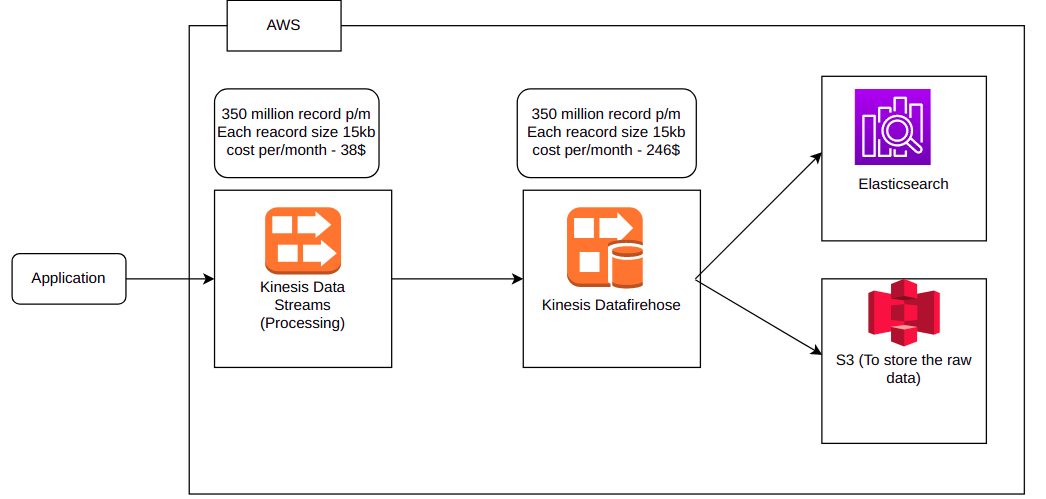
Apty wanted to do clickstream analysis (near real time) on AWS platform and for which they wanted to have small PoC of getting the data from API and ingesting into Elasticsearch through Kinesis Data Stream & Kinesis Firehose.

**Business Problem**

Apty wanted to do clickstream analysis (near real time) on a Web site in AWS Platform, clickstream analysis (also called clickstream analytics) is **the process of collecting, analysing and reporting aggregate data about which pages a website visitor visits -- and in what order**. The path the visitor takes though a website is called the clickstream.

**Solution Approach- Possible Approach**

Following Architecture been proposed



**Other Information mentioned in e-mail.**

 MRR: $ 818.86

  Cost link: <https://calculator.aws/#/estimate?id=f0ee7ddcbb63ea111712f5e94d11aff533b25a29>

**Assumptions**

**Data Stream**

~~-   4 million records~~

~~-  40 KB each record size~~

It was mentioned in previous discussion now it’s

350 million records per month

15 KB each record size

**Kinesis Data Analytics**

- 10 Apache Flink KPUs used per day

- 10 SQL KPUs per day

**Kinesis Data Firehose**

~~- 4 million records for data ingestion~~

~~- 40 KB each record size~~

350 million records per month

15 KB each record size

- Dynamic Partitioning enabled

**Lambda**

- 120 million requests per month

- Duration of each request 2ms

- Amount of memory allocated is 128 MB.

**Simple Storage Service (S3)**

- 5 TB of Standard storage.

**OpenSearch Service**

- 2 Data instances Nodes,

- t3 medium. search instance type,

- Pricing Model is OnDemand

- 50 GB Storage amount is for each instance

**Solution Approach considerations**

Following points need to be considered for Elastisearch (opensearch)

Please let us know the below configuration details regarding AWS Opensearch  
1. Deployment type (prod, dev/test)  
2. Version  
3. Data nodes-> Instance type, no. of nodes, EBS storage  
4. Network Configuration-> public access or Inside a VPC?  
5. Fine grained access control required or not?  
6. Encryption required?

7. Enablement of Ultra warm or cold storage required or not? If yes, will share the cost link.

KDS & Firehose Questions

8. How to integrate API with KDS

9. How many Firehose services need to be configured for Elastisearch and S3.

**As discussed with Customer on 13-Dec-21**

* They have already done PoC of workload few months back
* Only interested in AWS Credits
* Application integration would be taken by customer and only service provisioning needs to be done.