

CSAA Bootcamp – June 2019

Problem Scenario #2

Table of Contents

1	Document Definition	2
2	Problem Statement	2
3	Solution	2
3.1	Key Considerations	2
3.2	Technology Assumptions	2
3.3	AWS Services.....	3
3.4	Architecture Design	3
3.4.1	Solution 1	3
3.4.2	Solution 2	4

1 DOCUMENT DEFINITION

The purpose of this document is to present a solution to the Problem Scenario #2.

2 PROBLEM STATEMENT

PROBLEM SCENARIO



Problem Scenario 2

Overview

Most serverless architectures use API Gateway as the entry point for final users. This is since a RESTful API, agnostic of the implementation, is always a good service contract between our infrastructure and the final user. In some cases, this might impact cost but can be avoided.

In this scenario, API Gateway will execute a Lambda function that will pass the data to Kinesis Firehose and further on to Amazon S3. Here the cost comes from all these services.



The Challenge: Design 2 different approach to save cost.

3 SOLUTION

3.1 Key Considerations

The following are the key considerations for this solution:

1. Keep the core services in the current architecture to attain the same end result
2. Prioritize cost-saving solutions

3.2 Technology Assumptions

The following are the technology assumptions for this solution:

1. Like the current architecture, implementation-details will be excluded. Particularly, VPC and IGW settings.
2. Data processed from Kinesis Firehose will still be needed for retrieval at one point.
3. Data passed by the Client comes from a mobile application through AWS SDK.

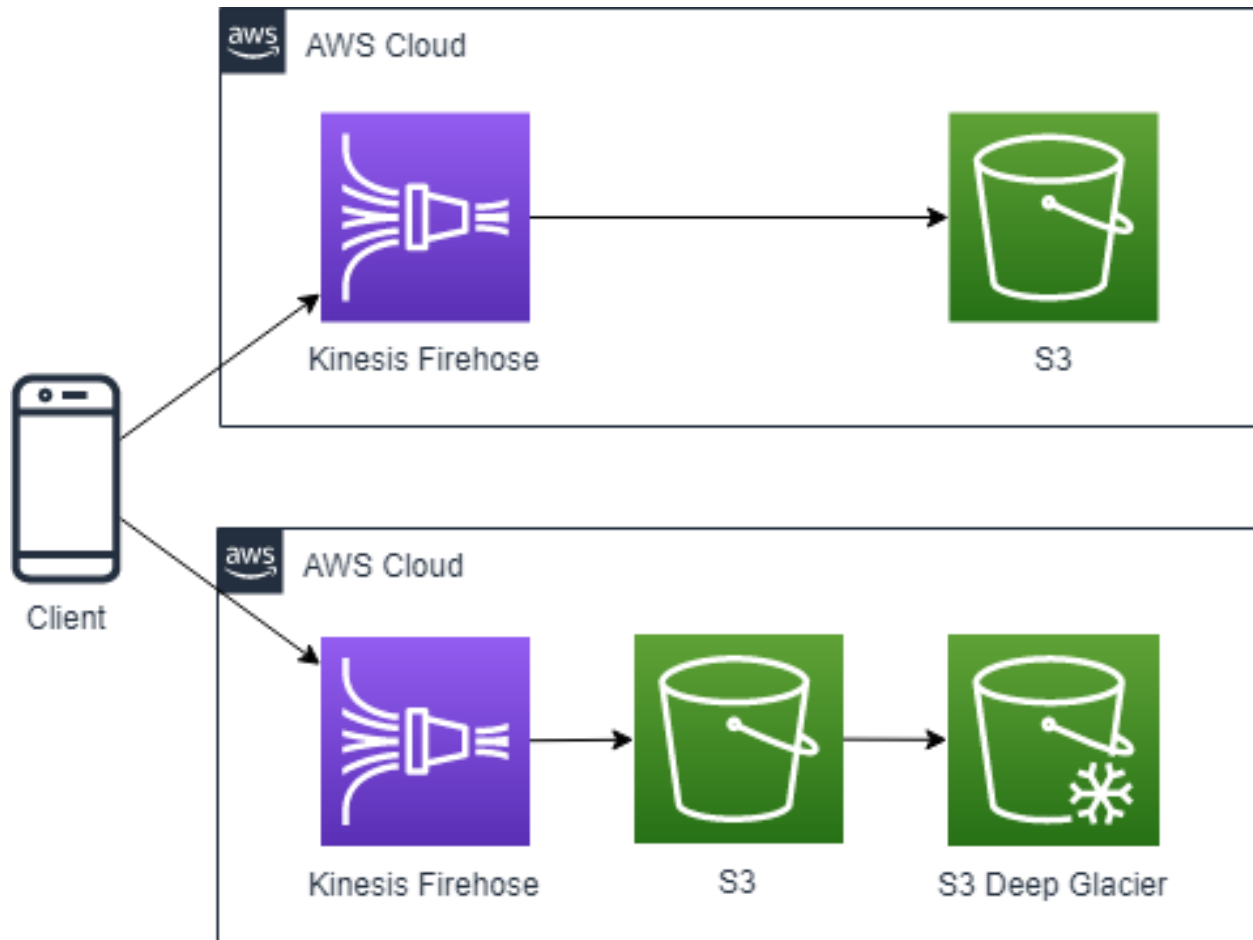
3.3 AWS Services

To achieve the considerations defined, the following AWS Services will be required:

1. Kinesis Firehose – processes streaming data
2. S3 – stores object-based data with retrieval in mind
3. S3 Deep Glacier – stores object-based data with archival in mind

3.4 Architecture Design

Below are two architecture designed to achieve the key considerations:



3.4.1 Solution 1

Data streamed from users will directly go to the Kinesis Firehose. In this case, Kinesis Firehose will process the data and forward the results to S3.

API Gateway and Lambda are removed since Kinesis Firehose is the only AWS service needed to process streaming data. Since the data from the client are already serialized through AWS SDK in the origin mobile application, there is no need to insist on having API Gateway. In addition, Kinesis Firehose already contains a

Lambda function by default to process the data. The external Lambda function in the current architecture is only utilized to “pass the data to Kinesis Firehose”.

S3 will still be retained to store the processed data from Kinesis Firehose for further processing which is out-of-scope for this scenario. Also, a lifecycle management will be configured to delete data after some interval (e.g., 6 months).

3.4.2 Solution 2

The same rationale is true for Solution 2 with an exception of an additional service, the S3 Deep Glacier. In this case, the lifecycle management will be reconfigured to send the data from S3 to S3 Deep Glacier after some interval (e.g., 6 months). An additional lifecycle management will be configured to delete the data in S3 Deep Glacier after some other interval (e.g., 1 year).