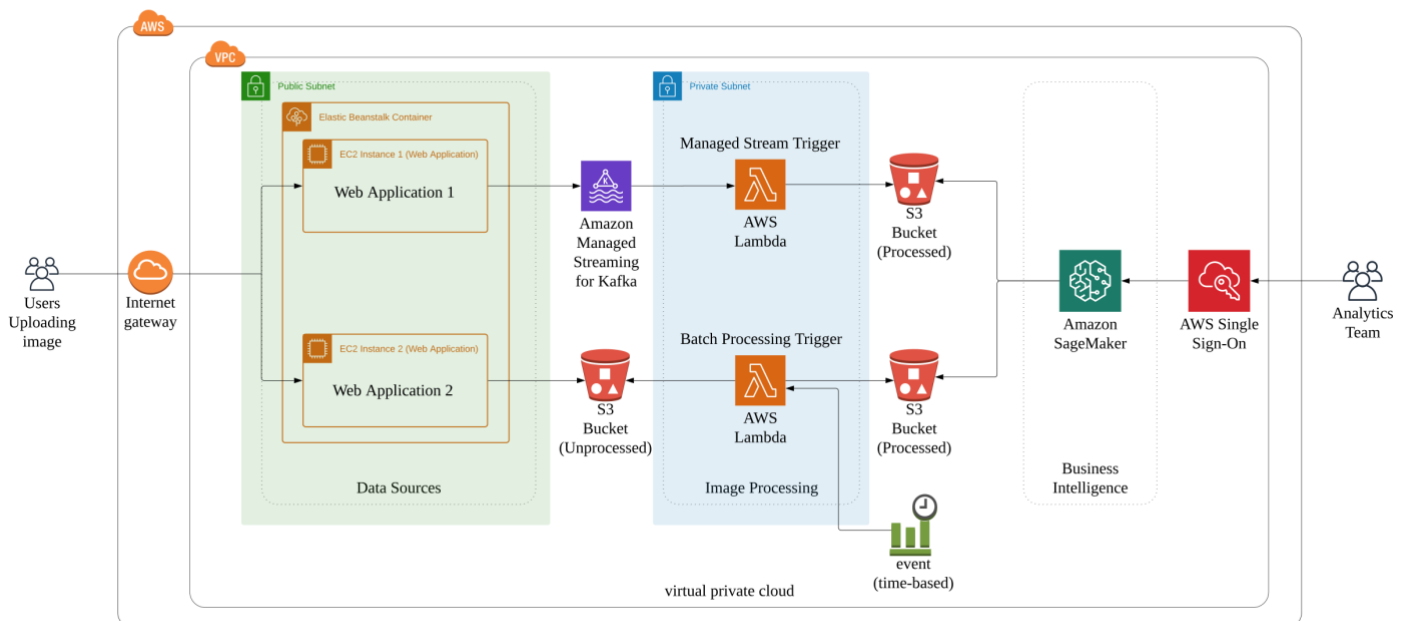


Section 3 System Design



The following assumptions are made about this solution:

1. Web application of the company are hosted using EC2 instance on AWS
2. Analytics team is using/knows how to use AWS SageMaker
3. Analytics team will be using IAM to access SageMaker to perform business intelligence task
4. System administrator is controlling the privilege and data access for user group of analytics team

Brief Write Up on System Architecture Design

General Overview

This solution leverages on services provided and offered by **Amazon Web Service (AWS)**. All services will be hosted within the **virtual private network VPC**. As this is a company whose main business is in processing images, the chosen storage would be the **AWS S3 bucket** as it is suitable for storing large amount of images and allows for extremely fast speeds of storage and retrieval.

Data Sources

In relation to the web application which is hosted on AWS EC2, **Elastic Beanstalk** is used to speed up and automate the process of the setup, configuration and Elastic load balancing. For web application 1 which streams the images using Kafka, the **Managed Streaming for Kafka** service is utilized, making it easier for developers and DevOps to run Apache Kafka applications on AWS. In relation to web application 2, once users uploads the image into the web application, the image would be stored in the S3 bucket.

Image Processing

The image will be processed using the code the company's software engineers have already written using AWS **Lambda** which is triggered via an event. For the Kafka stream, the event will be triggered by the stream while for the S3 bucket, it will be processed via a batch determined by the system admin using the **CloudWatch Event** where the admin is able to set a dedicated time to trigger the pre-processing of the image. Once pre-processed, it would then be stored in S3 for any amount of days determined by the system admin before the pre-processed image is removed.

Business Intelligence

Lastly, AWS **SageMaker** is also utilized for business intelligence as it allows for Jupyter notebooks which has an extremely low learning curve especially for analyst who wishes to do simple analytics. However, if required, SageMaker is able to perform extremely GPU intensive model training using more dedicated resources and along with its prebuilt models, the company is able to speed up the analytics process significantly. The analytics team access SageMaker via AWS single sign through an IAM provided by the admin. The team is able to build a image analytics model or conduct simple business intelligence using Jupyter notebook from SageMaker.